

# EXTREMALITY AND DYNAMICALLY DEFINED MEASURES, PART II: MEASURES FROM CONFORMAL DYNAMICAL SYSTEMS

TUSHAR DAS, LIOR FISHMAN, DAVID SIMMONS, AND MARIUSZ URBAŃSKI

**ABSTRACT.** We present a new method of proving the Diophantine extremality of various dynamically defined measures, vastly expanding the class of measures known to be extremal. This generalizes and improves the celebrated theorem of Kleinbock and Margulis ('98) resolving Sprindžuk's conjecture, as well as its extension by Kleinbock, Lindenstrauss, and Weiss ('04), hereafter abbreviated K LW. As applications we prove the extremality of all hyperbolic measures of smooth dynamical systems with sufficiently large Hausdorff dimension, and of the Patterson–Sullivan measures of all nonplanar geometrically finite groups. The key technical idea, which has led to a plethora of new applications, is a significant weakening of K LW's sufficient conditions for extremality.

In Part I, we introduce and develop a systematic account of two classes of measures, which we call *quasi-decaying* and *weakly quasi-decaying*. We prove that weak quasi-decay implies strong extremality in the matrix approximation framework, as well as proving the “inherited exponent of irrationality” version of this theorem.

In Part II, we establish sufficient conditions on various classes of conformal dynamical systems for their measures to be quasi-decaying. In particular, we prove the above-mentioned result about Patterson–Sullivan measures, and we show that Gibbs measures (including conformal measures) of nonplanar infinite iterated function systems (including those which do not satisfy the open set condition) and rational functions are quasi-decaying.

In subsequent parts, we will continue to exhibit numerous examples of quasi-decaying measures, in support of the thesis that “almost any measure from dynamics and/or fractal geometry is quasi-decaying”.

## CONTENTS

1. Introduction	2
1.1. Four conditions which imply strong extremality	3
1.2. Ahlfors regularity vs. exact dimensionality	4
1.3. Examples of friendly and absolutely friendly measures	5
1.4. Quasi-decay: examples and counterexamples	7
1.5. Main results: conformal examples of quasi-decaying measures	8
2. Proofs of Theorems 1.12 and 1.13	10
3. Patterson-Sullivan measures	11
3.1. Conformal and hyperbolic geometry	11
3.2. Geometrically finite groups	12
3.3. Patterson–Sullivan measures	13
3.4. From global decay to quasi-decay	13
3.5. Proof of global decay	14
3.6. Proof of friendliness	20
4. Gibbs measures of CIFSes	22
4.1. A general theorem	22
4.2. Definitions	25
4.3. Proof of Theorem 1.14	26
References	27

## 1. INTRODUCTION

In this series of papers we address a central problem in the flourishing area of metric Diophantine approximation on manifolds and measures: an attempt to exhibit a possibly widest natural class of sets and measures for which most points are not very well approximable by ones with rational coordinates.

Fix  $d \in \mathbb{N}$ . The quality of rational approximations to a vector  $\mathbf{x} \in \mathbb{R}^d$  can be measured by its *exponent of irrationality*, which is defined by the formula

$$\omega(\mathbf{x}) = \limsup_{\mathbf{p}/q \in \mathbb{Q}^d} \frac{-\log \|\mathbf{x} - \mathbf{p}/q\|}{\log(q)},$$

where the limsup is taken over any enumeration of  $\mathbb{Q}^d$ , and  $\|\cdot\|$  is any norm on  $\mathbb{R}^d$ . Another interesting quantity is the *exponent of multiplicative irrationality*, which is the number

$$\omega_{\times}(\mathbf{x}) = \limsup_{\mathbf{p}/q \in \mathbb{Q}^d} \frac{-\log \prod_{i=1}^d |x_i - p_i/q|}{\log(q)}.$$

It follows from a pigeonhole argument that  $\omega(\mathbf{x}) \geq 1 + 1/d$  and  $\omega_{\times}(\mathbf{x}) \geq d + 1$ . A vector  $\mathbf{x}$  is said to be *very well approximable* if  $\omega(\mathbf{x}) > 1 + 1/d$ , and *very well multiplicatively approximable* if  $\omega_{\times}(\mathbf{x}) > d + 1$ . We will denote the set of very well (multiplicatively) approximable vectors by  $\text{VW}(\text{M})A_d$ . It is well-known that  $\text{VWA}_d$  and  $\text{VWMA}_d$  are both Lebesgue nullsets of full Hausdorff dimension, and that  $\text{VWA}_d \subseteq \text{VWMA}_d$ .

A measure  $\mu$  on  $\mathbb{R}^d$  is *extremal* if  $\mu(\text{VWA}_d) = 0$ , and *strongly extremal* if  $\mu(\text{VWMA}_d) = 0$ . Extremality was first defined by V. G. Sprindžuk, who conjectured that the Lebesgue measure of any nondegenerate manifold is extremal. This conjecture was proven by D. Y. Kleinbock and G. A. Margulis [16], and later strengthened by D. Y. Kleinbock, E. Lindenstrauss, and B. Weiss (hereafter abbreviated “KLW”) in [15], who considered a class of measures which they called “friendly” and showed that these measures are strongly extremal. However, their definition is somewhat rigid and many interesting measures, in particular ones coming from dynamics, do not satisfy their condition. In this paper, we study a much larger class of measures, which we call *weakly quasi-decaying*, such that every weakly quasi-decaying measure is strongly extremal [Part I, Corollary 1.8]. This class includes a subclass of *quasi-decaying* measures, which are the analogue of KLW’s “absolutely friendly” measures.<sup>1</sup>

In the previous paper (Part I), we analyzed the basic properties of the quasi-decay and weak quasi-decay conditions, as well as proving that every weakly quasi-decaying measure is strongly extremal. In fact, we proved a more refined version of that theorem which we will not state here. We proved that every exact dimensional measure on  $\mathbb{R}^d$  whose dimension  $\delta$  satisfies  $\delta > d - 1$  is quasi-decaying. This elementary result already provides many dynamical examples of quasi-decaying measures; for example, a theorem of F. Hofbauer [13] states that any measure invariant under a piecewise smooth endomorphism of  $[0, 1]$  which has positive entropy is exact dimensional of positive dimension; since  $d = 1$  in this setup, this implies that such measures are quasi-decaying. (See §1.4 for more details.)

In the current paper (Part II) we continue with this theme, examining several classes of conformal dynamical systems to see under what circumstances the inequality  $\delta > d - 1$  can be relaxed to a “nonplanarity” assumption. The examples we will consider are intended to support the thesis that “any sufficiently non-pathological measure coming from dynamics or fractal geometry is quasi-decaying”. As a contrast, we also provide examples of non-extremal measures coming from dynamics, to gain a better insight as to which measures should be considered “pathological”.

In subsequent papers, we will continue to find examples of quasi-decaying measures, including random measures and measures invariant under diffeomorphisms.

**Convention 1.** The symbols  $\lesssim$ ,  $\gtrsim$ , and  $\asymp$  will denote coarse asymptotics; a subscript of  $+$  indicates that the asymptotic is additive, and a subscript of  $\times$  indicates that it is multiplicative. For example,  $A \lesssim_{\times, K} B$  means that there exists a constant  $C > 0$  (the *implied constant*), depending only on  $K$ , such that  $A \leq CB$ .

<sup>1</sup>The terminology “absolutely friendly” was not used by KLW and first appeared in [24]; however, several theorems about absolute friendliness had already appeared in [15] without using the terminology.

$A \lesssim_{+, \times} B$  means that there exist constants  $C_1, C_2 > 0$  so that  $A \leq C_1 B + C_2$ . In general, dependence of the implied constant(s) on universal objects will be omitted from the notation.

If  $\mu$  and  $\nu$  are measures, then  $\nu \lesssim_{\times} \mu$  means that there exists a constant  $C > 0$  such that  $\nu \leq C\mu$ .

**Convention 2.** In this paper, all measures and sets are assumed to be Borel, and measures are assumed to be locally finite.

**Convention 3.** For  $S \subseteq \mathbb{R}^d$  and  $\rho \geq 0$ ,  $\mathcal{N}(S, \rho) = \{x \in \mathbb{R}^d : d(x, S) \leq \rho\}$  is the closed  $\rho$ -thickening of  $S$ , and  $\mathcal{N}^\circ(S, \rho) = \{x \in \mathbb{R}^d : d(x, S) < \rho\}$  is the open  $\rho$ -thickening of  $S$ .

**Convention 4.**  $A \wedge B$  and  $A \vee B$  denote the minimum and maximum of  $A$  and  $B$ , respectively.

**Convention 5.** We use the Iverson bracket notation  $[\text{statement}] = \begin{cases} 1 & \text{statement true} \\ 0 & \text{statement false} \end{cases}$ .

**Convention 6.** The image of a measure  $\mu$  under a map  $f$  is denoted  $f_*[\mu] := \mu \circ f^{-1}$ .

**Convention 7.**  $\mathcal{H}$  denotes the collection of affine hyperplanes in  $\mathbb{R}^d$ .

**Convention 8.** The symbol  $\triangleleft$  will be used to indicate the end of a nested proof.

**Acknowledgements.** The first-named author was supported in part by a 2014-2015 Faculty Research Grant from the University of Wisconsin–La Crosse. The second-named author was supported in part by the Simons Foundation grant #245708. The third-named author was supported in part by the EPSRC Programme Grant EP/J018260/1. The fourth-named author was supported in part by the NSF grant DMS-1361677.

**1.1. Four conditions which imply strong extremality.** In [Part I, §1.1], we introduced the notions of quasi-decaying measures and weakly quasi-decaying measures, and compared them with KLW's notions of friendly and absolutely friendly measures. While we recall the definitions of the four classes of measures here, we refer to Part I for a more detailed discussion.

**Definition 1.1.** Let  $\mu$  be a measure on an open set  $U \subseteq \mathbb{R}^d$ , and let  $\text{Supp}(\mu)$  denote the topological support of  $\mu$ .

- $\mu$  is called *absolutely decaying* (resp. *decaying*) if there exist  $C_1, \alpha > 0$  such that for all  $\mathbf{x} \in \text{Supp}(\mu)$ ,  $0 < \rho \leq 1$ ,  $\beta > 0$ , and  $\mathcal{L} \in \mathcal{H}$ , if  $B = B(\mathbf{x}, \rho) \subseteq U$  then

$$(1.1) \quad \mu(\mathcal{N}^\circ(\mathcal{L}, \beta\rho) \cap B) \leq C_1 \beta^\alpha \mu(B) \quad (\text{absolutely decaying})$$

or

$$(1.2) \quad \mu(\mathcal{N}^\circ(\mathcal{L}, \beta\|d_{\mathcal{L}}\|_{\mu, B}) \cap B) \leq C_1 \beta^\alpha \mu(B) \quad (\text{decaying}),$$

respectively, where

$$\|d_{\mathcal{L}}\|_{\mu, B} := \sup\{d(\mathbf{y}, \mathcal{L}) : \mathbf{y} \in B \cap \text{Supp}(\mu)\}.$$

- $\mu$  is called *nonplanar* if  $\mu(\mathcal{L}) = 0$  for all  $\mathcal{L} \in \mathcal{H}$ . Note that every absolutely decaying measure is nonplanar. Moreover, the decaying and nonplanarity conditions can be combined notationally by using closed thickenings rather than open ones: a measure  $\mu$  is decaying and nonplanar if and only if there exist  $C_1, \alpha > 0$  such that for all  $\mathbf{x} \in \text{Supp}(\mu)$ ,  $0 < \rho \leq 1$ ,  $\beta > 0$ , and  $\mathcal{L} \in \mathcal{H}$ , if  $B = B(\mathbf{x}, \rho) \subseteq U$  then

$$(1.3) \quad \mu(\mathcal{N}(\mathcal{L}, \beta\|d_{\mathcal{L}}\|_{\mu, B}) \cap B) \leq C_1 \beta^\alpha \mu(B). \quad (\text{decaying and nonplanar})$$

- $\mu$  is called *Federer* (or *doubling*) if for some (equiv. for all)  $K > 1$ , there exists  $C_2 > 0$  such that for all  $\mathbf{x} \in \text{Supp}(\mu)$  and  $0 < \rho \leq 1$ , if  $B(\mathbf{x}, K\rho) \subseteq U$  then

$$(1.4) \quad \mu(B(\mathbf{x}, K\rho)) \leq C_2 \mu(B(\mathbf{x}, \rho)).$$

	Absolutely friendly	Friendly but not absolutely friendly	Not friendly
QD	<ul style="list-style-type: none"> <li>• Patterson–Sullivan measures of convex-cocompact groups</li> <li>• Gibbs measures of finite IFSeS and hyperbolic rational functions</li> </ul>	<ul style="list-style-type: none"> <li>• Patterson–Sullivan measures of geometrically finite groups which satisfy <math>k_{\min} &lt; d - 1</math></li> </ul>	<ul style="list-style-type: none"> <li>• Gibbs measures of nonplanar infinite IFSeS and rational functions</li> </ul>
WQD \setminus QD	Impossible	<ul style="list-style-type: none"> <li>• Lebesgue measures of nondegenerate manifolds</li> </ul>	<ul style="list-style-type: none"> <li>• Conformal measures of infinite IFSeS which have invariant spheres</li> <li>• Measures with finite Lyapunov exponent and zero entropy under the Gauss map</li> </ul>
E \setminus WQD	Impossible	Impossible	<ul style="list-style-type: none"> <li>• Generic invariant measures of hyperbolic toral endomorphisms</li> </ul>
Not E	Impossible	Impossible	<ul style="list-style-type: none"> <li>• Certain measures with infinite Lyapunov exponent under the Gauss map [11, Theorem 4.5]</li> </ul>

FIGURE 1. Some representative examples of dynamically defined measures compared on two axes: quasi-decay and friendliness. In the leftmost column we use the abbreviations QD = quasi-decaying, WQD = weakly quasi-decaying, and E = extremal.

If  $\mu$  is Federer, decaying, and nonplanar, then  $\mu$  is called *friendly*; if  $\mu$  is both absolutely decaying and Federer, then  $\mu$  is called *absolutely friendly*.<sup>2</sup> When the open set  $U$  is not explicitly mentioned, we assume that it is all of  $\mathbb{R}^d$ ; otherwise we say that  $\mu$  is absolutely decaying, friendly, etc. “relative to  $U$ ”.

**Definition 1.2.** Let  $\mu$  be a measure on  $\mathbb{R}^d$  and consider  $\mathbf{x} \in E \subseteq \mathbb{R}^d$ . We will say that  $\mu$  is *quasi-decaying* (resp. *weakly quasi-decaying*) at  $\mathbf{x}$  relative to  $E$  if for all  $\gamma > 0$ , there exist  $C_1, \alpha > 0$  such that for all  $0 < \rho \leq 1$ ,  $0 < \beta \leq \rho^\gamma$ , and  $\mathcal{L} \in \mathcal{H}$ , if  $B = B(\mathbf{x}, \rho)$  then

$$(1.5) \quad \mu(\mathcal{N}(\mathcal{L}, \beta\rho) \cap B \cap E) \leq C_1 \beta^\alpha \mu(B) \quad (\text{quasi-decaying})$$

or

$$(1.6) \quad \mu(\mathcal{N}(\mathcal{L}, \beta \|d_{\mathcal{L}}\|_{\mu, B}) \cap B \cap E) \leq C_1 \beta^\alpha \mu(B) \quad (\text{weakly quasi-decaying}),$$

respectively. We will say that  $\mu$  is *(weakly) quasi-decaying relative to  $E$*  if for  $\mu$ -a.e.  $\mathbf{x} \in E$ ,  $\mu$  is (weakly) quasi-decaying at  $\mathbf{x}$  relative to  $E$ . Finally, we will say that  $\mu$  is *(weakly) quasi-decaying* if there exists a sequence  $(E_n)_n$  such that  $\mu(\mathbb{R}^d \setminus \bigcup_n E_n) = 0$  and for each  $n$ ,  $\mu$  is (weakly) quasi-decaying relative to  $E_n$ .

We also recall that the following implications hold:

Absolutely friendly	$\Rightarrow$	Friendly
$\Downarrow$		$\Downarrow$
Quasi-decaying	$\Rightarrow$	Weakly quasi-decaying

Moreover, all four classes of measures are contained in the class of extremal measures. Now we are able to go farther, using examples from later in this paper as well as from the literature to show that these implications are all strict. See Figure 1 for more details.

**1.2. Ahlfors regularity vs. exact dimensionality.** One way of thinking about the difference between K LW’s conditions and our conditions is by comparing this difference with the difference between the classes of *Ahlfors regular* and *exact dimensional* measures, both of which are well-studied in dynamics. We recall their definitions:

**Definition.** A measure  $\mu$  on  $\mathbb{R}^d$  is called *Ahlfors  $\delta$ -regular* if there exists  $C > 0$  such that for every ball  $B(\mathbf{x}, \rho)$  with  $\mathbf{x} \in \text{Supp}(\mu)$  and  $0 < \rho \leq 1$ .

$$C^{-1} \rho^\delta \leq \mu(B(\mathbf{x}, \rho)) \leq C \rho^\delta.$$

<sup>2</sup>As K LW put it, the word “friendly” is “a somewhat fuzzy abbreviation of *Federer, nonplanar, and decaying*”.

The measure  $\mu$  is called *exact dimensional of dimension  $\delta$*  if for  $\mu$ -a.e.  $\mathbf{x} \in \mathbb{R}^d$ ,

$$(1.7) \quad \lim_{\rho \searrow 0} \frac{\log \mu(B(\mathbf{x}, \rho))}{\log \rho} = \delta.$$

The philosophical relations between Ahlfors regularity and exact dimensionality with absolute friendliness and quasi-decay, respectively, are:

$$(1.8) \quad \begin{aligned} \text{Ahlfors regular and “nonplanar”} &\Rightarrow \text{Absolutely friendly} \\ \text{Exact dimensional and “nonplanar”} &\Rightarrow \text{Quasi-decaying} \end{aligned}$$

Here “nonplanar” does not refer to nonplanarity as defined in Definition 1.1, but is rather something less precise (and stronger). This less precise definition should rule out examples like the Lebesgue measures of nondegenerate manifolds, since these are not quasi-decaying. One example of a “sufficient condition” for this imprecise notion of “nonplanarity” is simply the inequality  $\delta > d - 1$ , where  $\delta$  is the dimension of the measure in question. In particular, in this context the relations (1.8) are made precise by the following theorems:

**Theorem 1.3** ([17, Proposition 6.3]; cf. [24, 36]). *If  $\delta > d - 1$ , then every Ahlfors  $\delta$ -regular measure on  $\mathbb{R}^d$  is absolutely friendly.*

**Theorem 1.4** ([Part I, Theorem 1.5]). *If  $\delta > d - 1$ , then every exact dimensional measure on  $\mathbb{R}^d$  of dimension  $\delta$  is quasi-decaying.*

**1.3. Examples of friendly and absolutely friendly measures.** The two most canonical examples of friendly measures are the Lebesgue measure of a nondegenerate submanifold of  $\mathbb{R}^d$  and the image of an absolutely friendly measure under a nondegenerate embedding [15, Theorem 2.1]. Neither of these measures are absolutely friendly; other than Lebesgue measure, the most canonical example of an absolutely friendly measure is the Hausdorff measure of the limit set of a finite irreducible iterated function system satisfying the open set condition. The following is a slight refinement<sup>3</sup> of KLW’s original theorem regarding such measures:

**Theorem 1.5** ([7, Proposition 3.1], cf. [15, Theorem 2.3]). *Let  $\{u_1, \dots, u_m\}$  be a family of contracting similarities of  $\mathbb{R}^d$  satisfying the open set condition (see Definition 4.6), and let  $K$  be the limit set of this family (see Definition 4.7). If  $K$  is not contained in an affine hyperplane, then  $\mu = \mathcal{H}^\delta \llcorner K$  is absolutely friendly, where  $\delta = \dim_H(K)$ .*

Here and hereafter,  $\dim_H(S)$  denotes the Hausdorff dimension of a set  $S$ , and  $\mathcal{H}^s(S)$  denotes the  $s$ -dimensional Hausdorff measure of  $S$ .

Theorem 1.5 was generalized by the fourth-named author as follows:

**Theorem 1.6** ([35, Corollary 1.6]<sup>4</sup>). *Let  $\{u_1, \dots, u_m\}$  be a finite irreducible conformal iterated function system (CIFS) (see Definition 4.6) on  $\mathbb{R}^d$  satisfying the open set condition, and let  $K$  be the limit set of this family. Then  $\mu = \mathcal{H}^\delta \llcorner K$  is absolutely friendly, where  $\delta = \dim_H(K)$ .*<sup>5</sup>

<sup>3</sup>The refinement is to weaken the irreducibility hypothesis; in [15, Theorem 2.3] the assumption that  $K$  is not contained in any finite union of affine hyperplanes is required, but it turns out that this assumption is in fact equivalent to the more natural and formally weaker assumption that  $K$  is not contained in any affine hyperplane.

<sup>4</sup>Although the cited theorem includes the assumption  $d \geq 2$ , the case  $d = 1$  is true as a consequence of Theorem 1.3 and [19, Theorem 3.14].

<sup>5</sup>*Remark.* Although KLW write that Theorem 1.6 provides a partial answer to a conjecture of theirs [15, Conjecture 10.6 and “Added in proof” below], it actually provides a complete answer. Indeed, suppose that a CIFS satisfies the irreducibility assumption of KLW but not that of [35]. Then the limit set  $K$  of this CIFS is contained in a proper nondegenerate real-analytic submanifold  $M \subseteq \mathbb{R}^d$ , which if  $d \geq 3$  is a sphere. Let  $\phi : U \rightarrow M$  be a coordinate chart, where  $U \subseteq \mathbb{R}^{d-1}$ . If  $d \geq 3$ , assume that  $\phi$  is stereographic projection. Then  $\phi^{-1}(K)$  is the limit set of a CIFS which satisfies the irreducibility assumption of [35]. Thus by Theorem 1.6,  $\mathcal{H}^\delta \llcorner \phi^{-1}(K)$  is absolutely friendly, and by [15, Theorem 2.1(b)],  $\mathcal{H}^\delta \llcorner K \asymp_\times \phi_*[\mathcal{H}^\delta \llcorner \phi^{-1}(K)]$  is friendly and thus strongly extremal.

Note that the hypotheses of Theorems 1.5 and 1.6 imply (without the use of the irreducibility/nonplanarity hypothesis) that  $\mu$  is Ahlfors  $\delta$ -regular [19, Theorem 3.14], so the proposition provides an example of the philosophical interpretation (1.8) described above.

Actually, the main theorem of [35] concerned a class of measures more general than those of the form  $\mathcal{H}^\delta \llcorner K$ : the Gibbs measures of Hölder continuous potential functions. Such measures are not in general Federer unless a separation condition which is stronger than the open set condition is assumed.<sup>6</sup> This technicality caused a minor error in the statement of [35, Theorem 1.5];<sup>7</sup> the correct statement of that theorem should read as follows:

**Theorem 1.7** ([35, Theorem 1.5]). *Let  $\{u_1, \dots, u_m\}$  be a finite irreducible CIFS on  $\mathbb{R}^d$ . Let  $\phi : \{1, \dots, m\} \rightarrow \mathbb{R}$  be a Hölder continuous potential function, and let  $\mu_\phi$  be a Gibbs measure of  $\phi$ <sup>8</sup> (see Definition 4.8). If  $\mu_\phi$  is Federer (e.g. if  $\mu_\phi$  satisfies the strong separation condition, see Definition 4.6), then  $\mu_\phi$  is absolutely friendly.*

All three of these theorems concern measures related to conformal iterated function systems. There is a close relationship between CIFSes and two other classes of conformal dynamical systems: Kleinian groups and rational functions. The relationship between these last two classes is sometimes known as “Sullivan’s dictionary” and in [9, p.4], three of the authors proposed that the class of CIFSes should be added as a “third column” to Sullivan’s dictionary. Thus, we should expect to find analogues of Theorems 1.5-1.7 in the settings of Kleinian groups and rational functions. Indeed, the appropriate analogue in the setting of Kleinian groups was proven by B. O. Stratmann and the fourth-named author:

**Theorem 1.8** ([30]). *Let  $G$  be a convex-cocompact group of Möbius transformations of  $\mathbb{R}^d$  which does not preserve any generalized sphere (i.e. sphere or plane). Then the Patterson–Sullivan measure of  $G$  (see §3.3) is absolutely friendly.*

If this theorem is the analogue of Theorem 1.6, then the analogue of Theorem 1.7 should concern the Patterson densities of Hölder continuous Gibbs cocycles as defined in [22, p.3]. In fact, it is not hard to see that the proof of [30] generalizes to this setting; we omit the details, as this theorem is not relevant to the overall goal of this paper.

A more difficult variation of Theorem 1.8 involves moving to a different row of Sullivan’s dictionary (at least according to the table in [9, p.4]): the setting of geometrically finite groups with parabolic points. The reason for this difficulty is that such measures are in general not absolutely friendly, because if  $p$  is a parabolic point of a geometrically finite group  $G$  with limit set  $\Lambda$ , and  $\mathcal{L}$  is an affine hyperplane containing the tangent plane of  $\Lambda$  at  $p$ , then small neighborhoods of  $\mathcal{L}$  contain all the mass of small balls centered around  $p$ . This makes the Patterson–Sullivan measures of geometrically finite groups a good candidate for friendliness, since the quantity  $\|d_{\mathcal{L}}\|_{\mu, B}$  appearing in (1.2) is often much smaller than the radius of  $B$ . And indeed, we prove below:

**Theorem 1.9** (Proven in Section 3). *Let  $G$  be a geometrically finite group of Möbius transformations of  $\mathbb{R}^d$  (see §3.2) which does not preserve any generalized sphere. Then the Patterson–Sullivan measure  $\mu$  of  $G$  is friendly. Moreover,  $\mu$  is absolutely friendly if and only if every cusp of  $G$  has maximal rank (see §3.2).*

Moving to the third column of the dictionary, rational functions, it seems that the analogue of Theorems 1.5-1.8 has not been explicitly stated in the literature before, but can be proven using the techniques of [35] together with a rigidity result of W. Bergweiler, A. E. Eremenko, and S. J. van Strien [4, 10]:

**Theorem 1.10.** *Let  $T : \widehat{\mathbb{C}} \rightarrow \widehat{\mathbb{C}}$  be a hyperbolic (i.e. expansive on its Julia set) rational function, let  $\phi : \widehat{\mathbb{C}} \rightarrow \mathbb{R}$  be a Hölder continuous potential function, and let  $\mu_\phi$  be the corresponding Gibbs measure. If  $\text{Supp}(\mu_\phi)$  is not contained in any generalized sphere, then  $\mu_\phi$  is absolutely friendly.*

<sup>6</sup>For example, if  $u_a(x) = (x + a)/2$  and  $\phi_a(x) = \log(2^a/3)$  ( $a = 0, 1, x \in [0, 1]$ ), then  $\mu_\phi$  is the image under binary expansion of the Bernoulli measure on  $\{0, 1\}^{\mathbb{N}}$  corresponding to digit frequencies  $\text{freq}(0) = 1/3, \text{freq}(1) = 2/3$ ; it follows that for large values of  $n$  the intervals  $[1/2 - 1/2^n, 1/2]$  and  $[1/2, 1/2 + 1/2^n]$  have  $\mu_\phi$ -measures which are not comparable.

<sup>7</sup>Due to this error, Theorem 1.6 cannot be deduced as a corollary of Theorem 1.7, but since the measure in Theorem 1.6 is Federer by [19, Theorem 3.14], the proof of Theorem 1.7 actually proves Theorem 1.6 as well.

<sup>8</sup>Or more precisely, the image of a Gibbs measure of  $\phi$  under the coding map  $\pi : \{1, \dots, m\}^{\mathbb{N}} \rightarrow \mathbb{R}^d$ .

*Proof sketch.* Since  $T$  is hyperbolic, there is some metric on the Julia set  $J$  with respect to which  $T$  is distance expanding. Thus by [25, Theorem 4.5.2], the map  $T : J \rightarrow J$  has Markov partitions of arbitrarily small diameter. The inverse branches of  $T$  restricted to the elements of such a partition forms a graph directed system in the sense of [20], except that the cone condition is not necessarily satisfied. Now the Gibbs measure  $\mu_\phi$  is also a Gibbs measure of this graph directed system with respect to the potential function  $\phi$ . The proof of [35, Theorem 1.5] shows that if  $\mu_\phi$  is Federer and  $\text{Supp}(\mu_\phi)$  is not contained in any proper real-analytic submanifold of  $\widehat{\mathbb{C}}$ , then  $\mu_\phi$  is absolutely friendly. (The cone condition is not used in the proof of [35, Theorem 1.5], and the argument is easily extended from the realm of IFSeS to that of graph directed systems.) But the fact that  $\mu_\phi$  is a Gibbs measure for  $T$  implies that  $\mu_\phi$  is Federer (e.g. by modifying the proof of [26, Theorem A]). And if  $\text{Supp}(\mu_\phi)$  is not contained in any generalized sphere, then by [10, Corollary 1 and Theorem 2] (see also [4, Theorem 2]),  $\text{Supp}(\mu_\phi)$  is not contained in any proper real-analytic submanifold of  $\widehat{\mathbb{C}}$ . This completes the proof.  $\square$

It seems that this is a more or less complete list of those fractal measures which, before the invention of the quasi-decay condition, were known to be extremal. The results are restricted to one or possibly two rows of Sullivan's dictionary, corresponding to extremely rigid hypotheses (finiteness of alphabet, convex-cocompactness, hyperbolicity) on the class of dynamical systems considered. In the authors' view, these examples represent more or less the full scope of the friendliness and absolute friendliness conditions: rigid hypotheses were required because the friendliness and absolute friendliness hypotheses are themselves very rigid. By contrast, the quasi-decaying condition is quite flexible and exhibits a much broader range of examples; let us now get a feel for what these examples are.

**1.4. Quasi-decay: examples and counterexamples.** The simplest examples of measures which are quasi-decaying but not necessarily Federer come from applying Theorem 1.4 to measures invariant under one-dimensional dynamical systems. For these measures, exact dimensionality is known under fairly broad assumptions, and the dimension of exact dimensional measures can be computed directly:

**Theorem 1.11** ([13, Theorem 1]). *Let  $T : [0, 1] \rightarrow [0, 1]$  be a piecewise monotonic transformation whose derivative has bounded  $p$ -variation for some  $p > 0$ . Let  $\mu$  be a measure on  $[0, 1]$  which is ergodic and invariant with respect to  $T$ . Let  $h(\mu)$  and  $\chi(\mu)$  denote the entropy and Lyapunov exponent of  $\mu$ , respectively. If  $\chi(\mu) > 0$ , then  $\mu$  is exact dimensional of dimension*

$$\delta(\mu) = \frac{h(\mu)}{\chi(\mu)}.$$

Note that if  $h(\mu) > 0$ , then Ruelle's inequality<sup>9</sup> implies that  $\chi(\mu) > 0$ , so the above result applies and gives  $\delta_\mu > 0 = d - 1$ , so  $\mu$  is quasi-decaying.<sup>10</sup> On the other hand, in general there is no reason to believe that measures satisfying the hypotheses of Theorem 1.11 will be either Federer or decaying.

We should also note that Theorem 1.11 provides examples of dynamical measures which are not quasi-decaying as well. Namely, if  $\mu$  satisfies the hypotheses of Theorem 1.11 and  $h(\mu) = 0$ , then  $\mu$  is exact dimensional of dimension 0,<sup>11</sup> and so the following theorem shows that  $\mu$  is not quasi-decaying:

**Theorem 1.12** (Proven in Section 2). *Any exact dimensional measure of dimension 0 is not quasi-decaying.*

Actually, it is not surprising that such measures are not quasi-decaying, because the class of such measures includes some measures which are not extremal. Specifically, if  $T$  is the map  $x \mapsto nx$  for some  $n \geq 2$ , then the following theorem proves the existence and genericity of  $T$ -invariant measures which are not extremal:

<sup>9</sup>A proof valid in this context can be found in [3, Theorem 7.1].

<sup>10</sup>The inequality  $\chi(\mu) < \infty$  follows from the hypothesis that  $T'$  has bounded  $p$ -variation, which in particular implies that  $T'$  is bounded.

<sup>11</sup>In fact, the implication  $(h(\mu) = 0) \Rightarrow (\mu \text{ exact dimensional of dimension } 0)$  holds for higher-dimensional dynamical systems as well; see [2, comment following Proposition 2].

**Theorem 1.13** (Proven in Section 2). *Let  $T : X \rightarrow X$  be a hyperbolic toral endomorphism (cf. Definition 2.1), where  $X = \mathbb{R}^d / \mathbb{Z}^d$ . Let  $\mathbb{M}_T(X)$  be the space of  $T$ -invariant probability measures on  $X$ . Then the set of measures which give full measure to the Liouville points is comeager in  $\mathbb{M}_T(X)$ .*

In the one-dimensional case (i.e.  $T(x) = nx \bmod 1$  for some  $n \geq 2$ ), combining with Theorems 1.11 and 1.4 implies that the measures in Theorem 1.13 all have entropy zero with respect to  $T$ . Actually, this result, namely that generic invariant measures have entropy zero, holds more generally for all piecewise monotonic transformations of the interval [12, Theorem 2(vi)], as well as for all Axiom A diffeomorphisms [27, 6th theorem on p.101].

There are numerous other classes of measures coming from dynamics which are known to be exact dimensional. A notable example is the theorem of Barreira, Pesin, and Schmeling [2] to the effect that any measure ergodic, invariant, and hyperbolic with respect to a diffeomorphism is exact dimensional. Theorem 1.4 applies directly to those measures whose dimension is sufficiently large, but the question still remains about those measures whose dimension is not large enough. In Part III we will return to this question, answering it at least for the class of expanding toral endomorphisms with distinct eigenvalues, and possibly for a much larger class. But for now, we turn our attention to measures coming from conformal dynamical systems.

**1.5. Main results: conformal examples of quasi-decaying measures.** We now come to the main results of this paper, namely extensions of the results of §1.3 to much broader classes of conformal dynamical systems, replacing friendliness and absolute friendliness by the quasi-decay condition. Our first theorem and its corollary generalize Theorems 1.5-1.7. It will be proven via a more general theorem which will allow us to more easily deduce that certain measures are quasi-decaying in later papers in this series.

**Theorem 1.14** (Proven in Section 4). *Fix  $d \in \mathbb{N}$ , and let  $(u_a)_{a \in A}$  be an irreducible CIFS on  $\mathbb{R}^d$ . Let  $\phi : A^{\mathbb{N}} \rightarrow \mathbb{R}$  be a summable locally Hölder continuous potential function, let  $\mu_\phi$  be a Gibbs measure of  $\phi$ , and let  $\pi : A^{\mathbb{N}} \rightarrow \mathbb{R}^d$  be the coding map. Suppose that the Lyapunov exponent*

$$(1.9) \quad \chi_{\mu_\phi} := \int \log(1/|u'_{\omega_1}(\pi \circ \sigma(\omega))|) d\mu_\phi(\omega)$$

*is finite. Then  $\pi_*[\mu_\phi]$  is quasi-decaying.*

Letting  $\phi(\omega) = \delta \log(|u'_{\omega_1}(\pi \circ \sigma(\omega))|)$ , where  $\delta = \dim_H(K)$ , yields the following corollary:

**Corollary 1.15.** *Fix  $d \in \mathbb{N}$ , and let  $(u_a)_{a \in A}$  be a regular irreducible CIFS on  $\mathbb{R}^d$ . Let  $\mu$  be the conformal measure of  $(u_a)_{a \in A}$ , and let  $\pi : A^{\mathbb{N}} \rightarrow \mathbb{R}^d$  be the coding map. If the Lyapunov exponent of  $\mu$  is finite, then  $\pi_*[\mu]$  is quasi-decaying.*

The improvements on Theorems 1.6 and 1.7 are twofold:

- The CIFS can be infinite, as long as the Lyapunov exponent is finite.
- The open set condition is no longer needed.

Both of these improvements are quite significant. Without the open set condition it is hard to even calculate the dimension of the limit set of a CIFS; see e.g. [1, 23, 29]. Also, the geometry of infinite alphabet CIFSes can be much wilder than the geometry of finite CIFSes, whose limit sets are always Ahlfors regular (cf. [19, Theorem 3.14] versus [19, Lemma 4.12 - Theorem 4.16]). By contrast, the finite Lyapunov exponent assumption which replaces it is quite weak; for example, in the case of conformal measures it is implied by strong regularity (cf. [20, Definition 4.3.1]). It is also a necessary assumption, as demonstrated by certain IFSes related to continued fractions [11, Theorem 4.5].

A connection between the finite Lyapunov exponent condition and extremality also appeared in an earlier paper of three of the authors:

**Theorem 1.16** ([11, Theorem 2.1]). *If  $\mu$  is a probability measure on  $[0, 1] \setminus \mathbb{Q}$  invariant with finite Lyapunov exponent under the Gauss map, then  $\mu$  is extremal.*

This theorem is neither more nor less general than Theorem 1.14. It is obviously not more general, since it only applies to the Gauss map, and the conclusion that  $\mu$  is extremal is weaker than the conclusion of



Theorem 1.14 which states that the relevant measure is quasi-decaying. But it is also not less general, since it applies to all invariant measures with finite Lyapunov exponent, and not only those which are Gibbs measures of summable Hölder families of potential functions. The existence of invariant measures with finite Lyapunov exponent which are not quasi-decaying can be seen from Theorem 1.12, since the image of an entropy zero shift-invariant measure on  $\{1, 2\}^{\mathbb{N}}$  under the continued fractions / Gauss IFS coding map is such a measure. So Theorem 1.16 would become false if “extremal” were replaced by “quasi-decaying”. It would also become false if “Gauss map” were replaced by “shift map of a CIFS”, due to Theorem 1.13, which produces non-extremal measures invariant under the map  $x \mapsto 2x$ , which is the shift map of the binary IFS. This indicates that the phenomenon captured in Theorem 1.16 is really a “number-theoretic” phenomenon arising directly from the connection between the Gauss map and Diophantine approximation, rather than from some geometric intermediary.

Moving on to our next conformal setting, we consider the Patterson–Sullivan measures of geometrically finite Kleinian groups:

**Theorem 1.17** (Proven in Section 3). *Let  $G$  be a geometrically finite group of Möbius transformations of  $\mathbb{R}^d$  which does not preserve any generalized sphere. Then the Patterson–Sullivan measure  $\mu$  of  $G$  is quasi-decaying.*

At first, this theorem may not seem to give any new information beyond that provided in Theorem 1.9: that Patterson–Sullivan measures of irreducible geometrically finite groups are extremal. But actually, the quasi-decay condition implies more Diophantine properties than friendliness does, via [Part I, Theorem 1.3], which shows that the image of a quasi-decaying measure under a nondegenerate embedding is still extremal. Such a theorem cannot apply to friendly measures, since the Lebesgue measure of a nondegenerate manifold is friendly, but its image under a nondegenerate embedding may be contained in a rational hyperplane and therefore non-extremal. But Theorem 1.17 shows that such a fate cannot befall a Patterson–Sullivan measure. Two examples give this observation special significance: first of all, we may wish to consider the Patterson–Sullivan measure of a group of Möbius transformations of  $\mathbb{R}^d$  which preserves the unit sphere but not any smaller generalized sphere; this measure is absolutely continuous to the image under a nondegenerate embedding of a geometrically finite group acting irreducibly on  $\mathbb{R}^{d-1}$ , so Theorem 1.17 and [Part I, Theorem 1.3] imply that the measure is extremal. Second of all, if we consider  $\mathbb{R}^d$  as isomorphic with the space of  $M \times N$  matrices for some  $M, N \geq 2$  such that  $MN = d$ , then Theorem 1.17 and [Part I, Corollary 1.8] imply that the Patterson–Sullivan measure gives zero measure to the set of very well approximable  $M \times N$  matrices, but such a conclusion cannot be deduced from Theorem 1.9.

Also, treating Theorem 1.9 as though it “came first” is a bit silly in the sense that both theorems use the same main lemma (Lemma 3.8), and the argument from that lemma to Theorem 1.17 is significantly easier than the argument from that lemma to Theorem 1.9. In any case, it is interesting to have a natural example of a measure which is both quasi-decaying and friendly but not absolutely friendly.

Our last theorem relating quasi-decay to conformal dynamics is in the setting of rational functions. Its proof uses a recent theorem of M. Szostakiewicz, A. Zdunik, and the fourth-named author regarding “fine inducing”, as well as the rigidity result of W. Bergweiler, A. E. Eremenko, and S. J. van Strien mentioned earlier. We recall (cf. [14]) that if  $T : X \rightarrow X$  is a dynamical system, then a potential function  $\phi : X \rightarrow \mathbb{R}$  is called *hyperbolic* if there exists  $n \in \mathbb{N}$  such that  $\sup(S_n \phi) < P(T^n, S_n \phi)$ , where  $P(T, \phi)$  is the pressure of  $\phi$  with respect to  $T$ .

**Theorem 1.18.** *Let  $T : \widehat{\mathbb{C}} \rightarrow \widehat{\mathbb{C}}$  be a rational function, let  $\phi : \widehat{\mathbb{C}} \rightarrow \mathbb{R}$  be a Hölder continuous hyperbolic potential function, and let  $\mu_\phi$  be the Gibbs measure of  $(T, \phi)$ . If the Julia set of  $T$  is not contained in a generalized sphere, then  $\mu$  is quasi-decaying.*

*Proof.* Let  $(u_a)_{a \in A}$  be the CIFS described on [34, p.20]. By [34, Lemma 20], there exists a locally Hölder continuous potential function  $\bar{\phi}$  such that  $\mu_\phi$  is (the image under the coding map of) the Gibbs measure of  $\bar{\phi}$  with respect to  $(u_a)_{a \in A}$ . By [34, (22)], the function  $\bar{\phi}$  is summable, and by [34, Proposition 23], the measure  $\mu_\phi$  has finite Lyapunov exponent with respect to  $(u_a)_{a \in A}$ . So by Theorem 1.14, if  $(u_a)_{a \in A}$  acts irreducibly then  $\mu_\phi$  is quasi-decaying. But if  $(u_a)_{a \in A}$  acts reducibly, then  $\mu_\phi$  is supported on a proper real-analytic submanifold of  $\widehat{\mathbb{C}}$ , so the Julia set  $J = \text{Supp}(\mu_\phi)$  is contained in a proper real-analytic submanifold

of  $\widehat{\mathbb{C}}$ . So by [10, Corollary 1 and Theorem 2] (see also [4, Theorem 2]),  $J$  is contained in a generalized sphere.  $\square$

Comparing this theorem to Theorem 1.10, we see that we have replaced the hypothesis that the rational function  $T$  is hyperbolic with the hypothesis that the potential function  $\phi$  is hyperbolic. Although these conditions sound superficially similar due to the prolific use of the word “hyperbolic” in mathematical definitions, the latter is actually much weaker than the former, which essentially means that both critical points and parabolic points are irrelevant to the dynamics. By contrast, the hypothesis that  $\phi$  is hyperbolic does not actually place any restriction on the function  $T$  (given any rational function  $T$ , the function  $\phi \equiv 0$ , or more generally any function  $\phi$  satisfying  $\sup(\phi) - \inf(\phi) < \log \deg(T)$ , is hyperbolic), and in fact only fails when there is an equilibrium state whose entropy and Lyapunov exponent are both zero [14]. In particular, if  $T$  is a topological Collet–Eckmann map then every Hölder continuous potential function  $\phi$  is hyperbolic [14, Corollary 1.1].

**Outline of the paper.** In Section 2 we prove Theorems 1.12 and 1.13, thus providing some examples of measures which are not quasi-decaying. In Section 3 we prove Theorems 1.9 and 1.17, thus showing that Patterson–Sullivan measures of geometrically finite groups are both friendly and quasi-decaying. In Section 4 we prove Theorem 1.14, thus showing that the Gibbs measures of conformal iterated function systems are quasi-decaying.

**Remark.** We refer to [Part I, §3] for some preliminary results which will be used in the proofs, in particular the “quasi-Federer lemma” [Part I, Lemma 3.2] and the “Lebesgue differentiation theorem” [Part I, Theorem 3.6].

## 2. PROOFS OF THEOREMS 1.12 AND 1.13

*Proof of Theorem 1.12.* Let  $\mu$  be exact dimensional of dimension 0, and let  $E \subseteq \mathbb{R}^d$  be a set such that  $\mu(E) > 0$ . Then for a  $\mu$ -typical point  $\mathbf{x} \in E$ , by the exact dimensionality of  $\mu$  we have

$$\lim_{\rho \searrow 0} \frac{\log \mu(B(\mathbf{x}, \rho))}{\log(\rho)} = 0$$

while by the Lebesgue differentiation theorem [Part I, Theorem 3.6], [28, Theorem 9.1], we have

$$\lim_{\rho \searrow 0} \frac{\mu(B(\mathbf{x}, \rho) \cap E)}{\mu(B(\mathbf{x}, \rho))} = 1.$$

It follows that

$$\lim_{\rho \searrow 0} \log \frac{\mu(B(\mathbf{x}, \rho^\alpha) \cap E)}{\rho^\alpha \mu(B(\mathbf{x}, \rho))} = \infty \text{ for all } \alpha > 0,$$

which implies that  $\mu$  is not quasi-decaying at  $\mathbf{x}$  relative to  $E$ . Since  $\mathbf{x}$  was a  $\mu$ -typical point,  $\mu$  is not quasi-decaying relative to  $E$ ; since  $E$  was arbitrary subject to  $\mu(E) > 0$ ,  $\mu$  is not quasi-decaying.  $\square$

Before proving Theorem 1.13, we recall the definitions of some terms used in its statement.

**Definition 2.1.** Let  $X = \mathbb{R}^d / \mathbb{Z}^d$ . A *toral endomorphism* of  $X$  is a map  $T : X \rightarrow X$  of the form  $T([\mathbf{x}]) = [\mathbf{M}\mathbf{x}]$ , where  $\mathbf{M}$  is a  $d \times d$  matrix with integer entries. Here  $[\mathbf{x}]$  denotes the image of a point  $\mathbf{x} \in \mathbb{R}^d$  under the quotient map  $\mathbb{R}^d \rightarrow \mathbb{R}^d / \mathbb{Z}^d$ . The endomorphism  $T$  is called *hyperbolic* if the eigenvalues of  $\mathbf{M}$  all have modulus  $\neq 1$ .

*Proof of Theorem 1.13.* For each  $n \in \mathbb{N}$ , let

$$U_n = \bigcup_{\substack{\mathbf{p}/q \in \mathbb{Q} \\ q \geq n}} B\left(\frac{\mathbf{p}}{q}, \frac{1}{q^n}\right),$$

and let  $\mathcal{U}_n$  be the set of all measures  $\mu \in \mathbb{M}_T(X)$  such that  $\mu(U_n) > 1 - 2^{-n}$ . The sets  $U_n$  and  $\mathcal{U}_n$  are both open. The set  $G := \bigcap_n U_n$  is the set of Liouville points, i.e. points with infinite exponent of irrationality. Thus since every measure in  $\mathcal{G} := \bigcap_n \mathcal{U}_n$  gives full measure to  $G$ , it follows that no measure in  $\mathcal{G}$  is extremal.

To complete the proof, we need to show that  $\mathcal{G}$  is dense in  $\mathbb{M}_T(X)$ . Since  $\mathcal{G}$  is convex, it suffices to show that the closure of  $\mathcal{G}$  contains all ergodic measures in  $\mathbb{M}_T(X)$ . Indeed, let  $\mu \in \mathbb{M}_T(X)$  be an ergodic measure, and let  $\mathbf{x} \in X$  be a  $\mu$ -random point. Then by the ergodic theorem,

$$\mu_N := \frac{1}{N} \sum_{i=0}^{N-1} \delta_{T^i(\mathbf{x})} \rightarrow \mu,$$

where  $\delta$  denotes the Dirac delta. Fix  $\varepsilon > 0$ . Since  $T$  is hyperbolic, it has the specification property [5, (2.10) Theorem],<sup>12</sup> so there exists  $m = m_\varepsilon \in \mathbb{N}$  such that for all  $N \geq m$ , there exists  $\mathbf{y} = \mathbf{y}_{N,\varepsilon} \in X$  such that  $T^N(\mathbf{y}) = \mathbf{y}$  and

$$d(T^i(\mathbf{x}), T^i(\mathbf{y})) \leq \varepsilon \quad \forall i = 0, \dots, N - m - 1.$$

Next fix  $N \geq m$  and let

$$\nu_{N,\varepsilon} = \frac{1}{N} \sum_{i=0}^{N-1} \delta_{T^i(\mathbf{y})} \in \mathbb{M}_T(X).$$

If  $\mathbf{M}$  is the integer matrix representing  $T$ , then  $(\mathbf{M}^N - I)$  is a nonsingular integer matrix and  $(\mathbf{M}^N - I)\mathbf{y} \in \mathbb{Z}^d$ , so  $\mathbf{y} \in \mathbb{Q}^d$ . Thus  $\text{Supp}(\nu_{N,\varepsilon}) \subseteq \mathbb{Q}^d \subseteq G$ , so  $\nu_{N,\varepsilon} \in \mathcal{G}$ .

Let  $d$  be the co-Lipschitz distance on  $\mathbb{M}_T(X)$ , i.e.

$$d(\mu, \nu) = \sup_f \left| \int f \, d\mu - \int f \, d\nu \right|,$$

where the supremum is taken over all 1-Lipschitz functions  $f : X \rightarrow [-1, 1]$ . Then

$$\begin{aligned} d(\nu_{N,\varepsilon}, \mu) &\leq d(\nu_{N,\varepsilon}, \mu_N) + d(\mu_N, \mu) \\ &\leq d(\mu_N, \mu) + \frac{1}{N} \sum_{i=0}^{N-1} d(T^i(\mathbf{x}), T^i(\mathbf{y}_{N,\varepsilon})) \\ &\leq d(\mu_N, \mu) + \frac{1}{N} \left[ \sum_{i=0}^{N-m-1} \varepsilon + \sum_{i=N-m}^{N-1} 1 \right] \\ &\leq d(\mu_N, \mu) + \varepsilon + \frac{m_\varepsilon}{N} \xrightarrow{N \rightarrow \infty} \varepsilon \xrightarrow{\varepsilon \rightarrow 0} 0. \end{aligned}$$

Thus  $\mu$  is in the closure of the set  $\{\nu_{N,\varepsilon} : N \geq m, \varepsilon > 0\} \subseteq \mathcal{G}$ . □

### 3. PATTERSON-SULLIVAN MEASURES

In this section we will prove Theorems 1.17 and 1.9, namely that the Patterson–Sullivan measure of any irreducible geometrically finite Kleinian group is both quasi-decaying and friendly, but is absolutely friendly if and only if all cusps have full rank.

**3.1. Conformal and hyperbolic geometry.** We recall some preliminaries from conformal and hyperbolic geometry. Throughout this section,  $\mathbb{H} = \mathbb{H}^{d+1} = \{(x_0, \dots, x_d) \in \mathbb{R}^{d+1} : x_0 > 0\}$  and  $\mathbb{B} = \mathbb{B}^{d+1} = \{\mathbf{x} \in \mathbb{R}^{d+1} : \|\mathbf{x}\| < 1\}$  will denote the upper half-space and Poincaré ball models of hyperbolic geometry, respectively. Their boundaries are  $\partial\mathbb{H} = \mathbb{R}^d \cup \{\infty\}$  and  $\partial\mathbb{B} = S^d \subseteq \mathbb{R}^{d+1}$ , respectively. Isometries of  $\mathbb{H}$  and  $\mathbb{B}$  correspond to conformal isomorphisms of their boundaries. In particular,  $\mathbb{H}$  and  $\mathbb{B}$  are isometric via the *Cayley transform*  $\iota : \mathbb{H} \rightarrow \mathbb{B}$  defined by

$$\iota(\mathbf{x}) = 2 \frac{\mathbf{x} + \mathbf{e}_0}{\|\mathbf{x} + \mathbf{e}_0\|^2} - \mathbf{e}_0$$

and this induces (via the same formula) a conformal isomorphism between  $\partial\mathbb{H}$  and  $\partial\mathbb{B}$ , known as *stereographic projection*. The *spherical metric* on  $\partial\mathbb{H}$  is the pullback of the Euclidean metric on  $\partial\mathbb{B}$  under

<sup>12</sup>Although the result of [5] as stated only covers the case of invertible transformations, it is not too hard to prove the specification property for hyperbolic toral endomorphisms by direct calculation.

stereographic projection. We will denote it by  $D_s$ , while denoting the Euclidean metric on  $\partial\mathbb{H}$  by  $D_e$ . By contrast, we will denote the hyperbolic metric on  $\mathbb{H}$  by  $d_h$ , defined by the formula

$$\cosh d_h(\mathbf{x}, \mathbf{y}) = 1 + \frac{\|\mathbf{y} - \mathbf{x}\|^2}{2x_0y_0}$$

(cf. [8, (2.5.3)]). We will also use the subscripts s, e, and h in referring to constructions in which these metrics are used, e.g.  $B_h(\mathbf{x}, \rho)$ ,  $\mathcal{N}_e(\mathcal{L}, \beta\sigma)$ ,  $|g'(\mathbf{x})|_s$ , etc.

A *generalized sphere* in  $\partial\mathbb{H}$  is the pullback under stereographic projection of a set of the form  $\partial\mathbb{B} \cap \mathcal{A}$ , where  $\mathcal{A}$  is an affine subspace of  $\mathbb{R}^{d+1}$ . Equivalently, a generalized sphere is either an affine subspace of  $\partial\mathbb{H}$  (including  $\infty$  by convention) or the intersection of such a subspace with a sphere in  $\partial\mathbb{H}$ . The collection of all generalized spheres of  $\partial\mathbb{H}$  of codimension 1 (i.e. dimension  $d-1$ ) will be denoted  $\mathcal{S}$  (to contrast with  $\mathcal{H}$ , the collection of all hyperplanes). The collection  $\mathcal{S}$  is preserved under conformal transformations of  $\partial\mathbb{H}$ .

We denote the set of isometries of  $\mathbb{H}$  by  $\text{Isom}(\mathbb{H})$ . The following well-known results will be used frequently in the sequel:

**Theorem 3.1** (Geometric mean value theorem). *Fix  $g \in \text{Isom}(\mathbb{H})$ . Then for all  $\xi, \eta \in \partial\mathbb{H}$ , the formula*

$$(3.1) \quad \frac{D(g(\xi), g(\eta))}{D(\xi, \eta)} = (|g'(\xi)| \cdot |g'(\eta)|)^{1/2}$$

*holds both for the Euclidean metric and for the spherical metric.*

**Theorem 3.2** (Bounded distortion principle). *Fix  $g \in \text{Isom}(\mathbb{H})$  and let*

$$\|g\| = d_h(o, g(o)),$$

*where  $o = \mathbf{e}_0 \in \mathbb{H}$ . For all  $\xi \in \partial\mathbb{H}$  we have*

$$|g'(\xi)|_s \leq e^{\|g\|}.$$

*Moreover, if we fix  $\eta \in \partial\mathbb{H}$  and let*

$$\begin{aligned} d &= d_h(g^{-1}(o), [o, \eta]) \\ \rho &= D_s(\xi, \eta), \end{aligned}$$

*then*

$$|g'(\xi)| \gtrsim_{\times} e^{\|g\| - d} / (1 \vee (e^{\|g\|} \rho)).$$

*Here  $[o, \xi]$  denotes the geodesic ray connecting  $o$  and  $\xi$ .*

**3.2. Geometrically finite groups.** Fix  $G \leq \text{Isom}(\mathbb{H})$ . The *limit set* of  $G$ , denoted  $\Lambda$  is the collection of accumulation points in  $\partial\mathbb{H}$  of the set  $G(o)$ . The *convex hull* of the  $\Lambda$ , denoted  $\mathcal{C}_G$ , is the smallest convex subset of  $\mathbb{H}$  whose closure contains  $\Lambda$ , and is empty if  $\Lambda$  is a singleton. A *horoball* is a set of the form

$$H(\xi, t) = \{x \in \mathbb{H} : \mathcal{B}_\xi(o, x) > t\},$$

where  $\mathcal{B}_\xi$  denotes the *Busemann function*

$$\mathcal{B}_\xi(x, y) = \lim_{z \rightarrow \xi} [d_h(z, x) - d_h(z, y)].$$

The point  $\xi \in \partial\mathbb{H}$  is said to be the *center* of the horoball  $H(\xi, t)$ .

Next,  $G$  is said to be *geometrically finite* if there exist  $\sigma > 0$  and a disjoint  $G$ -invariant collection of horoballs  $\mathcal{H}$  such that

$$(3.2) \quad \mathcal{C}_G \subseteq G(B_h(o, \sigma)) \cup \bigcup(\mathcal{H}).$$

In this scenario, the collection  $\mathcal{H}$  can be written in the form

$$\mathcal{H} = \{H_\eta : \eta \in P\}$$

where  $P$  is the set of parabolic fixed points of  $G$  and for each  $\eta \in P$ ,  $H_\eta$  is a horoball centered at  $\eta$ . We will also need the following well-known results, in which  $G$  denotes a geometrically finite group:

**Theorem 3.3** (Cusp finiteness theorem). *There exists a finite set  $P_0$  such that  $P = G(P_0)$ .*

**Theorem 3.4** (Top representation theorem). *For all  $H \in \mathcal{H}$ , there exist  $\eta \in P_0$  and  $g \in G$  such that  $H = g(H_\eta)$  and*

$$d(g(o), [o, x]) \asymp_+ 0 \quad \forall x \in H.$$

**Theorem 3.5** (Boundedness of parabolic points). *For all  $\eta \in P$ , there exists a compact set  $\mathcal{D}_\eta \subseteq \Lambda$  not containing  $\eta$  such that*

$$\Lambda \setminus \{\eta\} = G_\eta(\mathcal{D}_\eta),$$

where  $G_\eta$  is the stabilizer of  $\eta$  in  $G$ .

**Theorem 3.6** (Translation planes). *For all  $\eta \in P$ , there exists a generalized sphere  $\mathcal{P}_\eta \subseteq \partial\mathbb{H}$  containing  $\eta$  such that  $G_\eta$  preserves  $\mathcal{P}_\eta$  and acts cocompactly on it.*

The dimension of  $\mathcal{P}_\eta$  is called the *rank* of  $\eta$ , and is denoted  $k_\eta$ .

**Theorem 3.7** (Inequality between rank and Poincaré exponent). *For all  $\eta \in P$ , we have  $\delta > k_\eta/2$ , where  $\delta$  is the Poincaré exponent*

$$(3.3) \quad \delta = \inf \left\{ s \geq 0 : \sum_{g \in G} e^{-s\|g\|} < \infty \right\}.$$

**3.3. Patterson–Sullivan measures.** The *Patterson–Sullivan measure* of  $G$ , denoted  $\mu$ , is the measure on  $\Lambda$ , unique up to a multiplicative factor, such that for all  $g \in G$  and  $E \subseteq \partial\mathbb{H}$ ,

$$(3.4) \quad \mu(g(E)) = \int_E |g'(\xi)|^\delta d\mu(\xi),$$

where  $\delta$  is as in (3.3). The existence of such a measure is due to S. J. Patterson and D. P. Sullivan [21, 32], and its uniqueness in the case of geometrically finite groups is due to Sullivan [33]. The support of the Patterson–Sullivan measure is the entire limit set, i.e.  $\text{Supp}(\mu) = \Lambda$ . Note that the Patterson–Sullivan measure is dependent on the choice of metric ( $D_s$  or  $D_e$ ), so there are really two Patterson–Sullivan measures  $\mu_s$  and  $\mu_e$ , related by the formula

$$\frac{d\mu_s}{d\mu_e}(\xi) = \left( \lim_{\eta \rightarrow \xi} \frac{D_s(\xi, \eta)}{D_e(\xi, \eta)} \right)^\delta.$$

Other than the transformation equation (3.4), the main facts we will need about the Patterson–Sullivan measure  $\mu$  are:

- $\mu$  is Federer; this follows from the global measure formula [33, §7], [31, Theorem 2].
- The following corollary of the logarithm law [31, Theorem 4]: for  $\mu$ -a.e.  $\xi \in \Lambda$ , we have

$$(3.5) \quad \lim_{t \rightarrow \infty} \frac{1}{t} d_h(\xi_t, G(o)) = 0,$$

where for each  $t \geq 0$ ,  $\xi_t$  denotes the unique point on the geodesic ray connecting  $o$  and  $\xi$  such that  $d_h(o, \xi_t) = t$ .

**3.4. From global decay to quasi-decay.** A group  $G \leq \text{Isom}(\mathbb{H})$  is said to be *irreducible* if  $G$  does not preserve any generalized sphere strictly contained in  $\partial\mathbb{H}$ . In the remainder of this section,  $G$  will denote an irreducible geometrically finite group,  $\mu$  its Patterson–Sullivan measure, etc. We will use the spherical metric as the default metric; in particular  $\mu$  denotes the Patterson–Sullivan measure with respect to the spherical metric. The key lemma which we will use to prove both quasi-decay and friendliness is the following:

**Lemma 3.8** (“Global decay”). *There exists  $\alpha > 0$  such that for all  $\beta > 0$  and  $\mathcal{L} \in \mathcal{S}$ , we have*

$$(3.6) \quad \mu(\mathcal{N}(\mathcal{L}, \beta)) \lesssim_\times \beta^\alpha.$$

Before proving this lemma, we use it to prove that  $\mu$  is quasi-decaying:

*Proof of Theorem 1.17 assuming Lemma 3.8.* Fix  $\xi \in \Lambda \setminus \{\infty\}$  satisfying (3.5), and we will show that  $\mu$  is quasi-decaying at  $\xi$  relative to  $\mathbb{R}^d$ . Since  $D_e \asymp_\times D_s$  in a small neighborhood of  $\xi$ , it is enough to show that, quantified appropriately, the equation (1.5) is satisfied when the metric is taken to be the spherical metric. Fix  $\gamma > 0$ ,  $0 < \rho \leq 1$ ,  $0 < \beta \leq \rho^\gamma$ , and  $\mathcal{L} \in \mathcal{H}$ . Replace  $\mathcal{L}$  by  $\mathcal{L} \cup \{\infty\}$ , so that  $\mathcal{L} \in \mathcal{S}$ .

Let  $t = -\log(\rho)$ , and fix  $g \in G$  such that  $d = d(\xi_t, g(o)) = d(\xi_t, G(o))$ . Then by Theorem 3.2,

$$(3.7) \quad e^{t-2d} \lesssim_\times |(g^{-1})'| \leq e^{t+d} \text{ on } B(\xi, \rho),$$

so

$$\begin{aligned} \mu(\mathcal{N}(\mathcal{L}, \beta\rho) \cap B(\xi, \rho)) &\lesssim_\times e^{-\delta(t-2d)} \mu(g^{-1}(\mathcal{N}(\mathcal{L}, \beta\rho))) && \text{(by (3.4))} \\ &\leq e^{-\delta(t-2d)} \mu(\mathcal{N}(g^{-1}(\mathcal{L}), e^{t+d}\beta\rho)) && \text{(by Theorem 3.1)} \\ &\lesssim_\times e^{-\delta(t-2d)} (e^{t+d}\beta\rho)^\alpha && \text{(by Lemma 3.8)} \\ &= \beta^\alpha \rho^\delta e^{-(\alpha+2\delta)d} \\ &\lesssim_\times \beta^\alpha \rho^{\delta-\alpha\gamma/4} && \text{(by (3.5))} \\ &\leq \beta^{\alpha/2} \rho^{\delta+\alpha\gamma/4}. && \text{(since } \beta \leq \rho^\gamma) \end{aligned}$$

A similar argument can be used to show that

$$(3.8) \quad \lim_{\rho \searrow 0} \frac{\log \mu(B(\xi, \rho))}{\log(\rho)} = \delta.$$

(Alternatively, (3.8) follows from (3.5) together with the global measure formula [31, Theorem 2], although this is overkill.) Thus

$$\rho^{\delta+\alpha\gamma/4} \lesssim_\times \mu(B(\xi, \rho)),$$

which completes the proof.  $\square$

**Remark 3.9.** The above proof is valid for any group  $G \leq \text{Isom}(\mathbb{H})$  (not necessarily geometrically finite) whose Patterson–Sullivan measure satisfies (3.5) and Lemma 3.8. It would be interesting to see whether there are geometrically infinite examples of groups with these properties.

**3.5. Proof of global decay.** We proceed to prove Lemma 3.8 via a series of reductions, finally reducing the question to one purely about Lebesgue measure.

**Lemma 3.10.** *There exist  $0 < \varepsilon, \lambda < 1$  such that for all  $0 < \beta \leq 1$  and  $\mathcal{L} \in \mathcal{S}$ , we have*

$$(3.9) \quad \mu(\mathcal{N}(\mathcal{L}, \varepsilon\beta)) \leq \lambda \mu(\mathcal{N}(\mathcal{L}, \beta)).$$

*Proof of Lemma 3.8 from Lemma 3.10.* Fix  $n \in \mathbb{N}$ . By applying (3.9) with  $\beta = 1, \varepsilon, \dots, \varepsilon^{n-1}$ , we have

$$\mu(\mathcal{N}(\mathcal{L}, \varepsilon^n)) \leq \lambda^n \mu(\partial\mathbb{H}) = \lambda^n.$$

Thus (3.6) holds with  $\alpha = \log_\varepsilon(\lambda)$ .  $\square$

**Lemma 3.11.** *For all  $\rho_0 > 0$ , there exists  $\varepsilon > 0$  with the following property: for all  $0 < \beta \leq 1$ ,  $\mathcal{L} \in \mathcal{S}$ , and  $\xi \in \Lambda \cap \mathcal{N}(\mathcal{L}, \varepsilon\beta)$ , there exists  $0 < \rho \leq \rho_0$  such that*

$$(3.10) \quad \mu(B(\xi, \rho) \cap \mathcal{N}(\mathcal{L}, \beta) \setminus \mathcal{N}(\mathcal{L}, \varepsilon\beta)) \asymp_\times \mu(B(\xi, \rho)).$$

*Proof of Lemma 3.10 from Lemma 3.11.* Let  $\rho_0 = 1$  and let  $\varepsilon$  be as in Lemma 3.11. Fix  $0 < \beta \leq 1$  and  $\mathcal{L} \in \mathcal{S}$ . For each  $\xi \in E := \text{Supp}(\mu) \cap \mathcal{N}(\mathcal{L}, \varepsilon\beta)$ , let  $0 < \rho_\xi \leq 1$  be chosen to satisfy (3.10). By the 4r-covering lemma [18, Theorem 8.1], the collection  $\{B(\xi, \rho_\xi) : \xi \in E\}$  has a disjoint subcollection

$\{B(\xi_i, \rho_i)\}_{i=1}^N$  such that the collection  $\{B(\xi_i, 4\rho_i)\}_{i=1}^N$  covers  $E$ . Then

$$\begin{aligned} \mu(\mathcal{N}(\mathcal{L}, \beta) \setminus \mathcal{N}(\mathcal{L}, \varepsilon\beta)) &\geq \sum_{i=1}^N \mu(B(\xi_i, \rho_i) \cap \mathcal{N}(\mathcal{L}, \beta) \setminus \mathcal{N}(\mathcal{L}, \varepsilon\beta)) \\ &\gtrsim_{\times} \sum_{i=1}^N \mu(B(\xi_i, \rho_i)) \quad (\text{Lemma 3.11}) \\ &\asymp_{\times} \sum_{i=1}^N \mu(B(\xi_i, 4\rho_i)) \quad (\text{since } \mu \text{ is Federer}) \\ &\geq \mu(E) = \mu(\mathcal{N}(\mathcal{L}, \varepsilon\beta)). \end{aligned}$$

After denoting the implied constant by  $C$  and letting  $\lambda = 1/(1 + 1/C) < 1$ , rearranging gives (3.9).  $\square$

**Reduction 3.12.** *In the proof of Lemma 3.11, we can without loss of generality assume that there exists  $\eta \in P_0$  (cf. Theorem 3.3) such that*

$$(3.11) \quad \varepsilon\beta \geq D^2(\xi, \eta)/C,$$

where  $C \geq 1$  is a constant possibly depending on  $\rho_0$ .

The basic idea is to “pull back” the entire picture via an isometry  $g \in G$ : to choose this isometry, we let  $x = \xi_t$  for an appropriate value of  $t$ , use (3.2) to find  $H \in \mathcal{H}$  such that  $x \in H$ , and then let  $g$  come from a top representation of  $H$  (cf. Theorem 3.4). After pulling back the picture, the new  $\beta$  is “on the large scale”, which translates quantitatively into the equation (3.11).

*Proof.* Suppose that for all  $C \geq 1$ , Lemma 3.11 holds in the special case (3.11). We let  $\rho_0, \varepsilon, \beta, \mathcal{L}, \xi, \rho$  denote the variables occurring in the version of Lemma 3.11 that we are trying to prove, while  $\widehat{\rho}_0, \widehat{\varepsilon}, \widehat{\beta}, \widehat{\mathcal{L}}, \widehat{\xi}, \widehat{\rho}$  denote the variables occurring in the version of Lemma 3.11 that we know. So fix  $\rho_0 > 0$ , and let  $\widehat{\rho}_0 = (1 \wedge \rho_0)/C_1$ , where  $C_1 \geq 1$  is large to be determined. Let  $0 < \widehat{\varepsilon} \leq 1/2$  be given, and let  $\varepsilon = \widehat{\varepsilon}/C_1$ , where  $C_2 \geq C_1$  is large to be determined, possibly depending on  $\widehat{\rho}_0$ . Fix  $0 < \beta \leq 1$ ,  $\mathcal{L} \in \mathcal{S}$ , and  $\xi \in \Lambda \cap \mathcal{N}(\mathcal{L}, \varepsilon\beta)$ .

Let  $t = -\log(C_2\varepsilon\beta)$  and  $x = \xi_t$ . Since  $x \in \mathcal{C}_G$ ,<sup>13</sup> by (3.2) we have either  $x \in g(B_h(\mathbf{0}, \sigma))$  for some  $g \in G$ , or else  $x \in H$  for some  $H \in \mathcal{H}$ . In the former case the reduction to the case  $\widehat{\varepsilon}\widehat{\beta} \geq 1/C$  follows from the comparison of the sets involved in (3.11) with their images under  $g^{-1}$ ; we omit the details as they are similar to what follows. (Alternatively, when  $G$  has at least one cusp, the former case can be reduced to the latter one by expanding the horoballs of  $\mathcal{H}$  so that they cover  $\mathcal{C}_G$  rather than being disjoint.)

So suppose that  $x \in H$  for some  $H \in \mathcal{H}$ . Let  $\eta \in P_0$  and  $g \in G$  be a top representation as in Theorem 3.4. Let

$$\widehat{\beta} = e^{\|g\|}\beta/C_1, \quad \widehat{\mathcal{L}} = g^{-1}(\mathcal{L}), \quad \text{and} \quad \widehat{\xi} = g^{-1}(\xi).$$

Since  $\xi \in \Lambda \cap \mathcal{N}(\mathcal{L}, \varepsilon\beta)$ , by Theorems 3.1 and 3.2,  $\widehat{\xi} \in \Lambda \cap \mathcal{N}(\widehat{\mathcal{L}}, \widehat{\varepsilon}\widehat{\beta})$ . On the other hand, since  $x \in H$  and  $g$  is a top representation, we have  $\mathcal{B}_{g(\eta)}(o, x) \gtrsim_+ \|g\|$  and thus

$$D^2(\xi, g(\eta)) \lesssim_{\times} e^{-t} e^{-\|g\|}.$$

Thus we get

$$D^2(\widehat{\xi}, \eta) \lesssim_{\times} e^{-t} e^{\|g\|} = C_2\varepsilon\beta e^{\|g\|} = C_2\widehat{\varepsilon}\widehat{\beta},$$

so by letting  $C = C_2 \cdot (\text{implied constant})$ , we get  $\widehat{\varepsilon}\widehat{\beta} \geq D^2(\widehat{\xi}, \eta)/C$ .

Suppose first that  $\widehat{\beta} \leq 1$ . Then the known special case of Lemma 3.11 applies; let  $0 < \widehat{\rho} \leq \widehat{\rho}_0$  be given as in that special case. Now since  $d(g(o), [o, \xi]) \leq d(g(o), [o, x]) \asymp_+ 0$ , we have  $D(g^{-1}(o), \widehat{\xi}) \gtrsim_{\times} 1$ . Let  $C_1$  be chosen so that  $1/C_1 \leq (1/2)D(g^{-1}(o), \widehat{\xi})$ ; then we have  $D(g^{-1}(o), \eta) \geq 1/C_1$  for all  $\eta \in B(\widehat{\xi}, 1/C_1)$ . So by Theorem 3.2, after possibly increasing  $C_1$  we have

$$(3.12) \quad e^{-\|g\|} \leq |g'| \leq C_1 e^{-\|g\|} \text{ on } B(\widehat{\xi}, 1/C_1).$$

<sup>13</sup>At the beginning of the proof, we can without loss of generality assume  $o \in \mathcal{C}_G$ , which then implies  $x \in \mathcal{C}_G$  at the current point of the proof.

Since  $\hat{\rho} \leq \hat{\rho}_0 \leq 1/C_1$ , the same holds on  $B(\hat{\xi}, \hat{\rho})$ . Thus, applying  $g$  to the known case of (3.10) and using Theorem 3.1, (3.4), and (3.12) gives

$$\mu(B(\xi, C_1 e^{-\|g\|} \hat{\rho}) \cap \mathcal{N}(\mathcal{L}, C_1 e^{-\|g\|} \hat{\beta}) \setminus \mathcal{N}(\mathcal{L}, e^{-\|g\|} \hat{\varepsilon} \hat{\beta})) \gtrsim_{\times} \mu(B(\xi, e^{-\|g\|} \hat{\rho})).$$

Letting  $\rho = C_1 e^{-\|g\|} \hat{\rho}$  and using the Federer condition gives the desired case of (3.10).

If  $\hat{\beta} > 1$ , then the above argument is still valid as long as there exists  $0 < \hat{\rho} \leq \hat{\rho}_0$  such that (3.10) holds. We claim that in fact, in this case (3.10) holds with  $\hat{\rho} = \hat{\rho}_0$ . Indeed, since  $\hat{\beta} > 1$ , we have  $\hat{\rho} \leq (1 - \hat{\varepsilon})\hat{\beta}$  assuming  $C_1 \geq 2$ , and thus since  $\hat{\xi} \in \mathcal{N}(\hat{\mathcal{L}}, \hat{\varepsilon}\hat{\beta})$ , we have

$$B(\hat{\xi}, \hat{\rho}) \subseteq \mathcal{N}(\hat{\mathcal{L}}, \hat{\beta})$$

and thus to complete the proof, it suffices to show that

$$(3.13) \quad \mu(\mathcal{N}(\hat{\mathcal{L}}, \hat{\varepsilon}\hat{\beta})) \leq (1/2)\mu(B(\hat{\xi}, \hat{\rho}_0)).$$

Now by a compactness argument, the right-hand side of (3.13) is bounded below by a constant depending only on  $\hat{\rho}_0$ . On the other hand, since  $x \in H$  and  $g$  is a top representation, we have  $t \gtrsim_+ \|g\|$  i.e.  $C_2 \varepsilon \beta \lesssim_{\times} e^{-\|g\|}$  i.e.  $\hat{\varepsilon}\hat{\beta} \lesssim_{\times} 1/C_2$  and thus the left-hand side of (3.13) is less than

$$f(C_2) = \sup_{\hat{\mathcal{L}} \in \mathcal{S}} \mu(\mathcal{N}(\hat{\mathcal{L}}, K/C_2)).$$

for some constant  $K$ . Now by a compactness argument and since  $G$  acts irreducibly, we have  $f(C_2) \rightarrow 0$  as  $C_2 \rightarrow \infty$ . So by choosing  $C_2$  sufficiently large, we can guarantee that the left-hand side of (3.13) is smaller than the right-hand side.  $\square$

The next reduction requires some motivation. Let  $\eta \in P_0$  be as in (3.11). Since  $P_0$  is finite, we can treat  $\eta$  as fixed. Let  $\mathcal{D} = \mathcal{D}_\eta$  and  $\mathcal{P} = \mathcal{P}_\eta$  be given by Theorems 3.5 and 3.6, respectively, let  $k = k_\eta = \dim(\mathcal{P})$ , and let  $H = G_\eta$  be the stabilizer of  $\eta$  in  $G$ . Without loss of generality we can assume that  $\mathcal{P} \setminus \{\eta\} \subseteq H(\mathcal{D})$ .

We proceed to approximate  $\mu$  by a smooth measure  $\lambda_\eta$  on  $\mathcal{P}$ . To motivate the choice of this measure, note that for each  $h \in H$ , the measure of the set  $h(\mathcal{D})$  can be computed using the transformation equation (3.4). Thus it will be useful if the new measure  $\lambda$  also satisfies (3.4), at least for  $h \in H$ . It is easy to come up with a formula for a measure on  $\mathcal{P}$  which satisfies (3.4) for an even larger class of Möbius transformations: namely, those  $h \in \text{Isom}(\mathbb{H})$  such that  $h(\eta) = \eta$  and  $h(\mathcal{P}) = \mathcal{P}$ . Precisely:

$$(3.14) \quad d\lambda_\eta(\xi) = D(\eta, \xi)^{2\delta-2k} d\lambda_{\mathcal{P}}(\xi).$$

Here  $\lambda_{\mathcal{P}}$  denotes the Hausdorff  $k$ -dimensional measure on  $\mathcal{P}$  with respect to the spherical metric  $D = D_s$ . Note that by Theorem 3.7, the singularity of  $\lambda_\eta$  at  $\eta$  is integrable, i.e.  $\lambda_\eta$  is a finite measure.

It turns out that when approximating  $\mu$  by  $\lambda_\eta$ , it is appropriate to replace every set of the form  $\mathcal{N}(\mathcal{L}, \beta)$  by a set of the form

$$\tilde{\mathcal{N}}(\mathcal{L}, \beta) := \mathcal{N}(\mathcal{L}, \beta) \cap B(\eta, \sqrt{\beta}).$$

We can now state the next lemma in the reduction:

**Lemma 3.13.** *For all  $\rho_0 > 0$ , there exists  $\varepsilon > 0$  such that for all  $0 < \beta \leq \rho_0^2/4$ ,  $\mathcal{L} \in \mathcal{S}$ , and  $\xi \in \mathcal{P} \cap \tilde{\mathcal{N}}(\mathcal{L}, \varepsilon\beta)$ , there exists  $\beta/2 \leq \rho \leq \rho_0$  such that*

$$(3.15) \quad \lambda_\eta(B(\xi, \rho) \cap \tilde{\mathcal{N}}(\mathcal{L}, \beta) \setminus \tilde{\mathcal{N}}(\mathcal{L}, \varepsilon\beta)) \asymp_{\times} \lambda_\eta(B(\xi, \rho)).$$

In the proof of Lemma 3.11 from Lemma 3.13, we will frequently use the asymptotic

$$(3.16) \quad |h'| \asymp_{\times} e^{-\|h\|} \asymp_{\times} D^2(\eta, h(\mathcal{D})) \text{ on } \mathcal{D} \ (h \in H),$$

which can be proven by conjugating  $\eta$  to  $\infty$  and then comparing the Euclidean and spherical metrics (noting that in the Euclidean metric, elements of  $H$  act as isometries). Note that by applying Theorem 3.1 to (3.16), we get

$$(3.17) \quad \text{Diam}(h(\mathcal{D})) \asymp_{\times} D^2(\eta, h(\mathcal{D})).$$



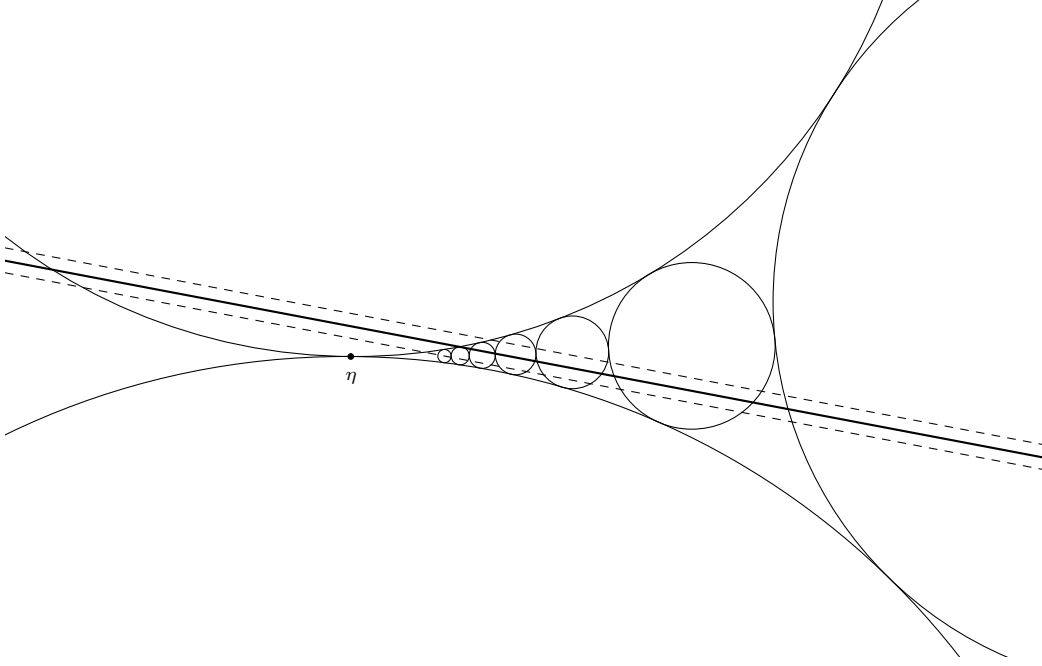


FIGURE 2. A schematic drawing of a hyperplane-neighborhood near a rank one cusp  $\eta$ . The balls represent sets  $h(\mathcal{D})$ ,  $h \in G_\eta$ . The measure of the hyperplane-neighborhood is estimated by considering its intersection with each of the sets  $h(\mathcal{D})$ .

*Proof of Lemma 3.11 using Lemma 3.13.* We let  $\rho_0, \varepsilon, \beta, \mathcal{L}, \xi, \rho$  denote the variables appearing in the desired Lemma 3.11, and we let  $\hat{\rho}_0, \hat{\varepsilon}, \hat{\beta}, \hat{\mathcal{L}}, \hat{\xi}, \hat{\rho}$  denote the variables appearing in the known Lemma 3.13. Fix  $\rho_0 > 0$ , choose  $\hat{\rho}_0 \leq \rho_0/3$  to be determined, let  $\hat{\varepsilon} > 0$  be given, and let  $\varepsilon = \hat{\varepsilon}/C^2 > 0$ , where  $C \geq 1$  is large to be determined. Fix  $0 < \beta \leq 1$ ,  $\mathcal{L} \in \mathcal{S}$ , and  $\xi \in \Lambda \cap \mathcal{N}(\mathcal{L}, \varepsilon\beta)$ , let  $\hat{\beta} = \beta/C$  and  $\hat{\mathcal{L}} = \mathcal{L}$ , and note that

$$\varepsilon\beta < \hat{\varepsilon}\hat{\beta} < \hat{\beta} < \beta.$$

By (3.17), we may choose  $\hat{\xi} \in \mathcal{P}$  so that  $D(\xi, \hat{\xi}) \lesssim_\times D^2(\eta, \xi)$ . Combining with (3.11) gives

$$D(\eta, \hat{\xi}) \lesssim_\times \sqrt{\varepsilon\beta}, \quad D(\hat{\xi}, \hat{\mathcal{L}}) \lesssim_\times \varepsilon\beta.$$

So by choosing  $C$  sufficiently large, we can guarantee that  $\hat{\xi} \in \tilde{\mathcal{N}}(\hat{\mathcal{L}}, \hat{\varepsilon}\hat{\beta})$ . Then we can let  $\hat{\beta}/2 \leq \hat{\rho} \leq \hat{\rho}_0$  be given as in Lemma 3.13. Let  $\rho = 3\hat{\rho} \leq 3\hat{\rho}_0 = \rho_0$ . To complete the proof we need to show:

$$(3.18) \quad \mu(B(\xi, \rho) \cap \mathcal{N}(\mathcal{L}, \beta) \setminus \mathcal{N}(\mathcal{L}, \varepsilon\beta)) \gtrsim_\times \lambda_\eta(B(\hat{\xi}, \hat{\rho}) \cap \tilde{\mathcal{N}}(\hat{\mathcal{L}}, \hat{\beta}) \setminus \tilde{\mathcal{N}}(\hat{\mathcal{L}}, \hat{\varepsilon}\hat{\beta}))$$

$$(3.19) \quad \mu(B(\xi, \rho)) \lesssim_\times \lambda_\eta(B(\hat{\xi}, 5\hat{\rho})).$$

((3.19) suffices since  $\lambda_\eta$  is Federer.) Fix  $h \in H$ , and we will demonstrate (3.18)-(3.19) via their intersections with  $h(\mathcal{D})$ , i.e. we will show that

$$(3.20) \quad \mu(h(\mathcal{D}) \cap B(\xi, \rho) \cap \mathcal{N}(\mathcal{L}, \beta) \setminus \mathcal{N}(\mathcal{L}, \varepsilon\beta)) \gtrsim_\times \lambda_\eta(h(\mathcal{D}) \cap B(\hat{\xi}, \hat{\rho}) \cap \tilde{\mathcal{N}}(\hat{\mathcal{L}}, \hat{\beta}) \setminus \tilde{\mathcal{N}}(\hat{\mathcal{L}}, \hat{\varepsilon}\hat{\beta}))$$

$$(3.21) \quad \mu(h(\mathcal{D}) \cap B(\xi, \rho)) \lesssim_\times \lambda_\eta(h(\mathcal{D}) \cap B(\hat{\xi}, 5\hat{\rho})).$$

The following consequence of (3.4) and (3.16) will be useful in proving both (3.20) and (3.21):

$$(3.22) \quad \mu(h(\mathcal{D})) \asymp_\times e^{-\delta\|h\|} \asymp_\times D^{2\delta}(\eta, h(\mathcal{D})) \asymp_\times \lambda_\eta(h(\mathcal{D})).$$

Also, by (3.11), if we choose  $C$  sufficiently large then

$$(3.23) \quad D(\xi, \hat{\xi}) \leq \hat{\rho}.$$

*Proof of (3.21).* To avoid trivialities, suppose that  $h(\mathcal{D}) \cap B(\xi, \rho) \neq \emptyset$ . Then by (3.11),

$$D(\eta, h(\mathcal{D})) \leq D(\eta, \xi) + \rho \lesssim_{\times} \sqrt{\varepsilon\beta} + \rho.$$

Applying (3.17) gives

$$\begin{aligned} \text{Diam}(h(\mathcal{D})) &\lesssim_{\times} \varepsilon\beta + \rho^2 = \widehat{\varepsilon}\widehat{\beta}/C + 4\widehat{\rho}^2 \\ &\leq (2\widehat{\varepsilon}/C + 4\widehat{\rho}_0)\widehat{\rho} \end{aligned} \quad (\text{since } \widehat{\beta}/2 \leq \widehat{\rho} \leq \widehat{\rho}_0)$$

and thus by choosing  $\widehat{\rho}_0$  small enough and  $C$  large enough, we get  $\text{Diam}(h(\mathcal{D})) \leq \widehat{\rho}$ . Combining with (3.23) gives  $h(\mathcal{D}) \subseteq B(\widehat{\xi}, 5\widehat{\rho})$ , and (3.22) completes the proof.  $\triangleleft$

*Proof of (3.20).* To avoid trivialities, suppose that

$$(3.24) \quad h(\mathcal{D}) \cap B(\widehat{\xi}, \widehat{\rho}) \cap \widetilde{\mathcal{N}}(\widehat{\mathcal{L}}, \widehat{\beta}) \setminus \widetilde{\mathcal{N}}(\widehat{\mathcal{L}}, \widehat{\varepsilon}\widehat{\beta}) \neq \emptyset.$$

By (3.23) we have  $B(\widehat{\xi}, \widehat{\rho}) \subseteq B(\xi, \rho)$  and thus  $h(\mathcal{D}) \cap B(\xi, \rho) \neq \emptyset$ , so the above argument shows that  $\text{Diam}(h(\mathcal{D})) \leq \widehat{\rho}$  and combining with (3.23) gives  $h(\mathcal{D}) \subseteq B(\xi, \rho)$ . On the other hand, since  $h(\mathcal{D}) \cap B(\eta, \widehat{\beta}^{1/2}) \neq \emptyset$ , we have  $D(\eta, h(\mathcal{D})) \leq \widehat{\beta}^{1/2}$  and thus by (3.17),  $\text{Diam}(h(\mathcal{D})) \lesssim_{\times} \widehat{\beta}$ . If  $C$  is sufficiently large, then this implies  $\text{Diam}(h(\mathcal{D})) \leq \beta/2$  and thus  $h(\mathcal{D}) \subseteq \mathcal{N}(\mathcal{L}, \beta)$ . So by (3.22), to complete the proof it is enough to show that

$$\mu(h(\mathcal{D}) \setminus \mathcal{N}(\mathcal{L}, \varepsilon\beta)) \asymp_{\times} \mu(h(\mathcal{D})).$$

Again to avoid trivialities, let us assume that

$$(3.25) \quad h(\mathcal{D}) \cap \mathcal{N}(\widehat{\mathcal{L}}, \varepsilon\beta) \neq \emptyset.$$

From (3.24), we have  $h(\mathcal{D}) \setminus \widetilde{\mathcal{N}}(\widehat{\mathcal{L}}, \widehat{\varepsilon}\widehat{\beta}) \neq \emptyset$ , so either

$$(3.26) \quad h(\mathcal{D}) \setminus \mathcal{N}(\widehat{\mathcal{L}}, \widehat{\varepsilon}\widehat{\beta}) \neq \emptyset$$

or

$$(3.27) \quad h(\mathcal{D}) \setminus B(\eta, (\widehat{\varepsilon}\widehat{\beta})^{1/2}) \neq \emptyset.$$

Combining (3.25), (3.26), and (3.17) gives  $e^{-\|h\|} \asymp_{\times} \text{Diam}(h(\mathcal{D})) \geq \widehat{\varepsilon}\widehat{\beta} - \varepsilon\beta$ , while combining (3.27) and (3.17) gives  $e^{-\|h\|} \asymp_{\times} D^2(\eta, h(\mathcal{D})) \gtrsim_{\times} \widehat{\varepsilon}\widehat{\beta}$ . Either way we get  $e^{-\|h\|} \gtrsim_{\times} \widehat{\varepsilon}\widehat{\beta}$ , and thus it suffices to show

$$\mu(h(\mathcal{D}) \setminus \mathcal{N}(\mathcal{L}, Ke^{-\|h\|}/C)) \asymp_{\times} \mu(h(\mathcal{D}))$$

where  $K > 1$  is a constant. By (3.16) and (3.4), it is enough to show that

$$(3.28) \quad \mu(\mathcal{D} \setminus \mathcal{N}(h^{-1}(\mathcal{L}), K/C)) \asymp_{\times} \mu(\mathcal{D}).$$

We proceed by contradiction; if no  $C$  exists satisfying (3.28) (for all  $\mathcal{L} \in \mathcal{S}$  and  $h \in H$ ), then a compactness argument proves the existence of  $\mathcal{L}_0 \in \mathcal{S}$  such that  $\mu(\mathcal{D} \setminus \mathcal{L}_0) = 0$ . A zooming argument shows that we may take  $\mathcal{L}_0$  so that  $\mu(\Lambda \setminus \mathcal{L}_0) = 0$ , i.e.  $\Lambda \subseteq \mathcal{L}_0$ . But this contradicts the assumption that  $G$  is irreducible.  $\triangleleft$

This completes the proof of Lemma 3.11 modulo Lemma 3.13.  $\square$

**Reduction 3.14.** *In the proof of Lemma 3.13, we can without loss of generality assume that  $\mathcal{P}$  is a plane (rather than a sphere) and that  $\eta = \mathbf{0}$ , and we can work in the Euclidean metric rather than the spherical metric.*

*Proof.* The first reduction follows by applying a fixed conjugation in which we move  $\eta$  to  $\mathbf{0}$  and some other point of  $\mathcal{P}$  to  $\infty$ . The second reduction follows from choosing  $\rho_0$  small enough so that  $D_e \asymp_{\times} D_s$  on  $B_s(\eta, \rho_0)$ , and modifying the constants  $\varepsilon, \beta, \rho$  appropriately.  $\square$

We are now ready to finish the proof of Theorem 1.17 by proving the Euclidean version of Lemma 3.13:

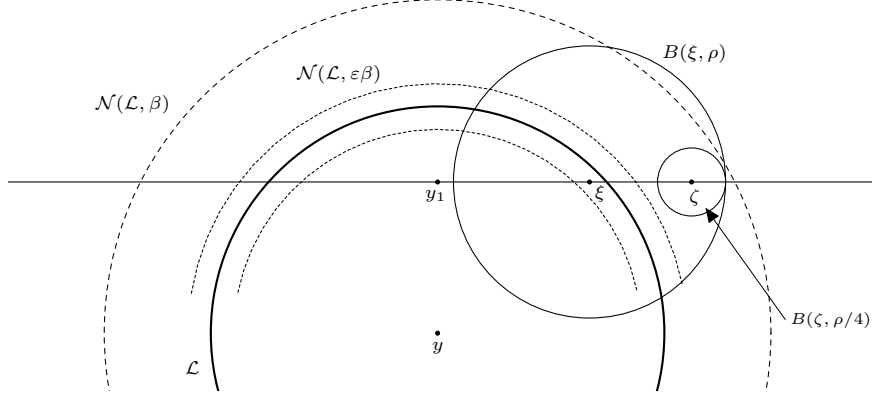


FIGURE 3. Various entities appearing in the proof of Lemma 3.13, illustrating the inclusion (3.29) in the case  $\rho = D(\xi, \mathcal{P} \setminus \mathcal{N}(\mathcal{L}, \beta))$ .

*Proof of Lemma 3.13.* Let  $\rho = D(\xi, \mathcal{P} \setminus \tilde{\mathcal{N}}(\mathcal{L}, \beta))$ . Then

$$\rho \geq D(\tilde{\mathcal{N}}(\mathcal{L}, \varepsilon\beta), \mathcal{P} \setminus \tilde{\mathcal{N}}(\mathcal{L}, \beta)) \geq (\beta - \varepsilon\beta) \wedge (\sqrt{\beta} - \sqrt{\varepsilon\beta}) \geq \beta/2$$

and on the other hand,  $\rho \leq \text{Diam}(\tilde{\mathcal{N}}(\mathcal{L}, \beta)) \leq 2\sqrt{\beta} \leq \rho_0$ . By construction,  $B(\xi, \rho) \cap \mathcal{P} \subseteq \tilde{\mathcal{N}}(\mathcal{L}, \beta)$ . Since the measure  $\lambda_\eta$  is Federer, to prove (3.15) it suffices to show that there exists a ball  $B(\zeta, \rho/4)$  such that  $\zeta \in \mathcal{P}$  and

$$(3.29) \quad B(\zeta, \rho/4) \cap \mathcal{P} \subseteq B(\xi, \rho) \setminus \tilde{\mathcal{N}}(\mathcal{L}, \varepsilon\beta).$$

We consider two cases:

- Suppose  $\rho = D(\xi, \mathcal{P} \setminus B(\mathbf{0}, \sqrt{\beta})) = \sqrt{\beta} - \|\xi\|$ . Let  $\mathbf{v} \in \mathcal{P}$  be a unit vector in the direction of  $\xi$  (in an arbitrary direction if  $\xi = \mathbf{0}$ ) and let  $\zeta = \xi + 3\rho\mathbf{v}/4$ . Then for all  $\mathbf{z} \in B(\zeta, \rho/4)$ ,

$$\|\mathbf{z}\| \geq \|\zeta\| - \rho/4 \geq \|\xi\| + \rho/2 \geq \sqrt{\beta}/2 \geq \sqrt{\varepsilon\beta}$$

assuming  $\varepsilon \leq 1/4$ . Thus  $B(\zeta, \rho/4) \cap B(\mathbf{0}, \sqrt{\varepsilon\beta}) = \emptyset$ , demonstrating (3.29).

- Suppose  $\rho = D(\xi, \mathcal{P} \setminus \mathcal{N}(\mathcal{L}, \beta))$ . Then  $\mathcal{P} \not\subseteq \mathcal{L}$ , since otherwise  $\rho = \infty$  which contradicts the definition of  $\rho$ . Suppose that  $\mathcal{L}$  is a sphere (the case where  $\mathcal{L}$  is a hyperplane is easier and will be omitted), and write  $\mathcal{L} = \{\mathbf{x} : \|\mathbf{y} - \mathbf{x}\| = k\}$  for some  $\mathbf{y} \in \mathbb{R}^d$  and  $k > 0$ . Write  $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2$  with  $\mathbf{y}_1 \in \mathcal{P}$  and  $\mathbf{y}_2 \in \mathcal{P}^\perp$ . Let  $\mathbf{v} \in \mathcal{P}$  be a unit vector in the direction of  $\xi - \mathbf{y}_1$  (in an arbitrary direction if  $\xi = \mathbf{y}_1$ ), and let  $\zeta = \xi + 3\rho\mathbf{v}/4$ . Then for all  $\mathbf{z} \in B(\zeta, \rho/4) \cap \mathcal{P}$ ,

$$D(\mathbf{z}, \mathcal{L}) = \|\mathbf{z} - \mathbf{y}\| - k \geq \sqrt{(\|\xi - \mathbf{y}_1\| + \rho/2)^2 + \|\mathbf{y}_2\|^2} - k.$$

We aim to show that the right hand side exceeds  $\varepsilon\beta$ . Indeed, since  $\xi \in \tilde{\mathcal{N}}(\mathcal{L}, \varepsilon\beta)$  we have

$$\varepsilon\beta \geq D(\xi, \mathcal{L}) \geq k - \|\xi - \mathbf{y}\| = k - \sqrt{\|\xi - \mathbf{y}_1\|^2 + \|\mathbf{y}_2\|^2}$$

and since  $\xi + \rho\mathbf{v} \in B(\xi, \rho) \subseteq \partial\tilde{\mathcal{N}}(\mathcal{L}, \beta)$ , we have

$$\beta = D(\xi + \rho\mathbf{v}, \mathcal{L}) = \|\xi + \rho\mathbf{v}\| - k = \sqrt{(\|\xi - \mathbf{y}_1\| + \rho)^2 + \|\mathbf{y}_2\|^2} - k,$$

i.e.

$$\begin{aligned} \sqrt{\|\xi - \mathbf{y}_1\|^2 + \|\mathbf{y}_2\|^2} - k &\geq -\varepsilon\beta \\ \sqrt{(\|\xi - \mathbf{y}_1\| + \rho)^2 + \|\mathbf{y}_2\|^2} - k &\geq \beta \end{aligned}$$

which implies

$$\sqrt{(\|\xi - \mathbf{y}_1\| + \rho/2)^2 + \|\mathbf{y}_2\|^2} - k \geq (1/4)\beta + (3/4)(-\varepsilon\beta).^{14}$$

Assuming  $\varepsilon \leq 1/7$ , this gives  $D(\mathbf{z}, \mathcal{L}) \geq \varepsilon\beta$  and thus (3.29) holds.  $\square$

**3.6. Proof of friendliness.** Like the proof of Lemma 3.8, the proof of Theorem 1.9 proceeds via a series of reductions, in which the final question is about Lebesgue measure and does not depend on the measure. Since many of the arguments are similar to those in the proof of Lemma 3.8, we will not provide the full details of the reductions.

We will actually prove a result which is slightly stronger than friendliness, namely “friendliness to spheres”:

**Theorem 3.15** (Friendliness to spheres). *There exists  $\alpha > 0$  such that for all  $\xi \in \Lambda$ ,  $0 < \rho \leq 1$ ,  $\beta > 0$ , and  $\mathcal{L} \in \mathcal{S}$ , if  $B = B(\xi, \rho)$  then*

$$\mu(\mathcal{N}(\mathcal{L}, \beta \|d_{\mathcal{L}}\|_{\mu, B}) \cap B) \lesssim_{\times} \beta^{\alpha} \mu(B),$$

and if  $k_{\min} = d$  then

$$\mu(\mathcal{N}(\mathcal{L}, \beta \rho) \cap B) \lesssim_{\times} \beta^{\alpha} \mu(B).$$

Note that Theorem 3.15 implies the hard direction of Theorem 1.9. The easy direction is proven as follows: Suppose that  $k_{\eta} < d$  for some  $\eta$ , and we will show that  $\mu$  is not absolutely decaying. By Theorems 3.5 and 3.6, there exist two generalized spheres  $\mathcal{L}_1, \mathcal{L}_2 \in \mathcal{S}$  which are tangent at  $\eta$  such that  $\Lambda$  is contained in the region between  $\mathcal{L}_1$  and  $\mathcal{L}_2$ . Let  $\mathcal{L}$  be the hyperplane tangent to both  $\mathcal{L}_1$  and  $\mathcal{L}_2$  at  $\eta$ . Then for all  $\rho > 0$ ,

$$\Lambda \cap B(\eta, \rho) \subseteq \mathcal{N}(\mathcal{L}, C\rho^2),$$

where  $C > 0$  is a large constant. This implies that  $\mu$  is not absolutely decaying (even stronger,  $\Lambda$  is not hyperplane diffuse in the sense of [6, Definition 4.2]).

We now proceed with the proof of Theorem 3.15.

**Reduction 3.16.** *In the proof of Theorem 3.15, we can without loss of generality assume that*

$$(3.30) \quad \rho \gtrsim_{\times} D^2(\xi, \eta)$$

for some  $\eta \in P_0$ , where  $P_0$  is as in Theorem 3.3.

*Proof.* Similar to the proof of Reduction 3.12.  $\square$

As in the proof of Lemma 3.8, fix  $\eta \in P_0$ , and let  $\mathcal{D} = \mathcal{D}_{\eta}$ ,  $\mathcal{P} = \mathcal{P}_{\eta}$ ,  $k = k_{\eta} = \dim(\mathcal{P})$ ,  $H = G_{\eta}$ , and  $\lambda_{\eta}$  be as before. We will prove the reduction of Theorem 3.15 to the following lemma in a manner analogous to the reduction of Lemma 3.11 to Lemma 3.13:

**Lemma 3.17.** *There exists  $\alpha_0 > 0$  such that for all  $0 < \alpha \leq \alpha_0$ ,  $\xi \in \mathcal{P}$ ,  $0 < \rho \leq 1$ ,  $\beta > 0$ , and  $\mathcal{L} \in \mathcal{S}$ , if  $B = B(\xi, \rho)$  and*

$$\sigma_1 = \max_B D^2(\eta, \cdot), \quad \sigma_2 = \|d_{\mathcal{L}}\|_{B \cap \mathcal{P}},$$

then either

$$(3.31) \quad \int_{B \cap \mathcal{P}} \left( \frac{\beta(\sigma_1 \vee \sigma_2)}{D^2(\eta, \mathbf{y})} \right)^{\alpha} d\lambda_{\eta}(\mathbf{y}) \lesssim_{\times} \beta^{\alpha/2} \lambda_{\eta}(B)$$

or

$$(3.32) \quad \lambda_{\eta}(\mathcal{N}(\mathcal{L}, K\sigma_1 \vee \beta\sigma_2) \cap B) \lesssim_{\times} \beta^{\alpha/2} \lambda_{\eta}(B),$$

where  $K > 0$  is the implied constant of (3.17).

<sup>14</sup>The general inequality

$$\sqrt{(a + \delta x)^2 + b^2} \geq \delta^2 \sqrt{(a + x)^2 + b^2} + (1 - \delta^2) \sqrt{a^2 + b^2}$$

( $a, b, x \geq 0$ ,  $0 \leq \delta \leq 1$ ) can be verified by first checking the inequality

$$(a + \delta x)^2 + b^2 \geq \delta^2[(a + x)^2 + b^2] + (1 - \delta^2)[a^2 + b^2]$$

and then using the downward convexity of the square root function.

*Proof of Theorem 3.15 using Lemma 3.17.* We let  $\alpha, \xi, \rho, \beta, \mathcal{L}$  denote the variables appearing in the desired Theorem 3.15, and we let  $\hat{\alpha}_0, \hat{\alpha}, \hat{\xi}, \hat{\rho}, \hat{\beta}, \hat{\mathcal{L}}$  denote the variables appearing in the known Lemma 3.17. Let  $\hat{\alpha}_0 > 0$  be given, let  $\tilde{\alpha} > 0$  be as in Lemma 3.8, and let  $\alpha = \hat{\alpha} = \hat{\alpha}_0 \wedge \tilde{\alpha} > 0$ . Fix  $\xi \in \Lambda$ ,  $D^2(\xi, \eta) \lesssim_{\times} \rho \leq 1$ ,  $\beta > 0$ , and  $\mathcal{L} \in \mathcal{S}$ , and let  $B = B(\xi, \rho)$ . By (3.17), we may choose  $\hat{\xi} \in \mathcal{P}$  so that  $D(\xi, \hat{\xi}) \lesssim_{\times} D^2(\eta, \xi) \lesssim_{\times} \rho$ . Fix  $C \geq 1$  large to be determined, and let  $\hat{\rho} = C\rho > 0$ , so that  $B(\xi, \rho) \subseteq \hat{B} := B(\hat{\xi}, \hat{\rho})$  assuming  $C$  is large enough. Let  $\hat{\beta} = C\beta$  and  $\hat{\mathcal{L}} = \mathcal{L}$ . Finally, let

$$\sigma = \begin{cases} \|d_{\mathcal{L}}\|_{\mu, B} & k_{\min} < d \\ \rho & k_{\min} = d \end{cases}, \quad \hat{\sigma}_1 = \max_{\hat{B}} D^2(\eta, \cdot), \quad \hat{\sigma}_2 = \|d_{\mathcal{L}}\|_{\hat{B} \cap \mathcal{P}}.$$

Note that by (3.17), we have

$$(3.33) \quad \sigma \lesssim_{\times} \hat{\sigma}_1 \vee \hat{\sigma}_2.$$

(When  $k_{\min} = d$ , (3.33) follows from the asymptotic  $\hat{\sigma}_2 \asymp_{\times} \hat{\rho} \geq \rho$ .) Since  $\mu$  is Federer, to complete the proof it suffices to show that

$$(3.34) \quad \mu(\mathcal{N}(\mathcal{L}, \beta\sigma) \cap B) \lesssim_{\times} \int_{\hat{B} \cap \mathcal{P}} \left( \frac{\hat{\beta}(\hat{\sigma}_1 \vee \hat{\sigma}_2)}{D^2(\eta, \mathbf{y})} \right)^{2\alpha} d\lambda_{\eta}(\mathbf{y})$$

$$(3.35) \quad \mu(\mathcal{N}(\mathcal{L}, \beta\sigma) \cap B) \lesssim_{\times} \lambda_{\eta}(\mathcal{N}(\hat{\mathcal{L}}, K\hat{\sigma}_1 \vee \hat{\beta}\hat{\sigma}_2) \cap \hat{B})$$

$$(3.36) \quad \mu(B(\xi, 2\hat{\rho})) \gtrsim_{\times} \lambda_{\eta}(\hat{B})$$

Fix  $h \in H$ , and we will prove (3.34)-(3.36) via their intersections with  $h(\mathcal{D})$ . The proof of (3.36) is similar to the proof of (3.19). Suppose  $h(\mathcal{D}) \cap \mathcal{N}(\mathcal{L}, \beta\sigma) \cap B \neq \emptyset$ , so that by (3.17) and (3.30), we have  $h(\mathcal{D}) \subseteq \mathcal{N}(\hat{\mathcal{L}}, K\hat{\sigma}_1 \vee \hat{\beta}\hat{\sigma}_2) \cap \hat{B}$ . Then (3.35) follows from (3.22), and (3.34) is reduced to

$$(3.37) \quad \mu(\mathcal{N}(\mathcal{L}, \beta\sigma) \cap h(\mathcal{D})) \lesssim_{\times} (e^{\|h\|} \hat{\beta}(\hat{\sigma}_1 \vee \hat{\sigma}_2))^{2\alpha} \lambda_{\eta}(h(\mathcal{D})).$$

Applying (3.16), (3.1), and (3.4) reduces us to proving

$$(3.38) \quad \mu(\mathcal{N}(h^{-1}(\mathcal{L}), K_2 e^{\|h\|} \beta\sigma) \cap \mathcal{D}) \lesssim_{\times} (e^{\|h\|} \hat{\beta}(\hat{\sigma}_1 \vee \hat{\sigma}_2))^{2\alpha} \lambda_{\eta}(\mathcal{D}),$$

where  $K_2 > 1$  is the implied constant of (3.16). But this follows from (3.6) and (3.33).  $\square$

**Reduction 3.18.** *In the proof of Lemma 3.17, we can without loss of generality assume that  $\mathcal{P}$  is a plane (rather than a sphere) and that  $\eta = \mathbf{0}$ , and we can work in the Euclidean metric rather than the spherical metric.*

*Proof.* Similar to the proof of Reduction 3.14.  $\square$

*Proof of Lemma 3.17.* It suffices to prove

$$(3.39) \quad \int_{B \cap \mathcal{P}} \left( \frac{\sigma_1}{D^2(\mathbf{0}, \mathbf{y})} \right)^{\alpha} d\lambda_{\mathbf{0}}(\mathbf{y}) \lesssim_{\times} \lambda_{\mathbf{0}}(B)$$

$$(3.40) \quad \lambda_{\mathbf{0}}(\mathcal{N}(\mathcal{L}, 2K\beta^{1/2}\sigma_2) \cap B) \lesssim_{\times} (2K\beta^{1/2})^{\alpha} \lambda_{\mathbf{0}}(B)$$

since the former implies (3.31) if  $\sigma_1 \geq \beta^{1/2}\sigma_2$ , and the latter implies (3.32) if  $\sigma_1 \leq \beta^{1/2}\sigma_2$ . Using the asymptotic

$$\lambda_{\mathbf{0}} \upharpoonright B(\mathbf{0}, 1) \asymp_{\times} \sum_{n \in \mathbb{N}} 2^{n(2k-2\delta)} \lambda_{\mathcal{P}}^{(e)} \upharpoonright B(\mathbf{0}, 2^{-n}) \setminus B(\mathbf{0}, 2^{-(n+1)}),$$

one can show that it suffices to consider the case where  $D(\mathbf{0}, B) \geq \rho$ . Here  $\lambda_{\mathcal{P}}^{(e)}$  is the Euclidean Lebesgue measure on  $\mathcal{P}$ . Now if  $D(\mathbf{0}, B) \geq \rho$ , then  $D(\mathbf{0}, \mathbf{y}) \asymp_{\times} D(\mathbf{0}, B) \asymp_{\times} \sigma_1^{1/2}$  for all  $\mathbf{y} \in B$ , which implies (3.39).  $\square$

**Lemma 3.19.** *If  $\lambda$  is Lebesgue measure on a  $k$ -dimensional subspace  $\mathcal{P} \subseteq \mathbb{R}^d$ ,  $\mathcal{L} \in \mathcal{S}$ , and  $B(\xi, \rho) \subseteq \mathcal{P}$ , then*

$$(3.41) \quad \lambda(\mathcal{N}(\mathcal{L}, \beta\sigma) \cap B(\xi, \rho)) \lesssim_{\times} \beta^{\alpha} \lambda(B(\xi, \rho)),$$

where  $\sigma = \|d_{\mathcal{L}}\|_{B(\xi, \rho) \cap \mathcal{P}}$ , and  $\alpha > 0$  is a uniform constant.

*Proof.* Without loss of generality suppose that  $\|\xi\|, \rho \leq 1$ . Then we can use the spherical metric instead of the Euclidean metric. Using the spherical metric, (3.41) is just the assertion that the image of Lebesgue measure under stereographic projection onto the sphere is friendly. But this follows from [15, Theorem 2.1].  $\square$

#### 4. GIBBS MEASURES OF CIFSES

We will prove Theorem 1.14 by first proving a general theorem about measures which come from “coding maps” (Theorem 4.3). This general theorem will be useful in later papers in this series, where we will use it to deduce that certain random measures are quasi-decaying.

**4.1. A general theorem.** Before stating the general theorem, we state a lemma which essentially says that when a measure is defined as the image of another measure, then the quasi-decaying condition can be checked on the level of the original measure rather than on the level of the image measure:

**Lemma 4.1.** *Let  $(X, \mu)$  be a measure space, and let  $\pi : X \rightarrow \mathbb{R}^d$  be a measurable map. Suppose that there exists a sequence of sets  $(E_n)_1^\infty$  in  $X$  such that  $\mu(X \setminus \bigcup_n E_n) = 0$  and for all  $n \in \mathbb{N}$ , for  $\mu$ -a.e.  $x \in E_n$ , for all  $\gamma > 0$ , there exist  $C_1, \alpha > 0$  such that for all  $0 < \rho \leq 1$ ,  $0 < \beta \leq \rho^\gamma$ , and  $\mathcal{L} \in \mathcal{H}$ , if  $B = B(\pi(x), \rho)$  then*

$$(4.1) \quad \mu(\pi^{-1}(\mathcal{N}(\mathcal{L}, \beta\rho) \cap B) \cap E_n) \leq C_1 \beta^\alpha \mu(\pi^{-1}(B)).$$

Then  $\bar{\mu} := \pi_*[\mu]$  is quasi-decaying.

*Proof.* Without loss of generality assume that the sequence  $(E_n)_1^\infty$  is increasing. For each  $n$ , define the measure  $\nu_n$  on  $\mathbb{R}^d$  via the formula

$$\nu_n(S) = \mu(\pi^{-1}(S) \cap E_n),$$

i.e.  $\nu_n = \pi_*[\mu \upharpoonright E_n]$ . Then  $\nu_n \nearrow \bar{\mu}$ . Let  $f_n$  denote the Radon–Nikodym derivative  $\frac{d\nu_n}{d\bar{\mu}}$ , so that  $f_n \nearrow 1$   $\bar{\mu}$ -a.e. We may choose  $f_n$  so that  $f_n = 0$  on  $\mathbb{R}^d \setminus \pi(E_n)$ . Finally, let  $F_n = \{\mathbf{y} \in \mathbb{R}^d : f_n(\mathbf{y}) \geq 1/2\}$ , so that  $\bar{\mu}(\mathbb{R}^d \setminus \bigcup_n F_n) = 0$ . For each  $n$ , we can see that  $\bar{\mu}$  is quasi-decaying relative to  $F_n$  as follows: given  $\mathbf{y} \in F_n$  and  $\gamma > 0$ , choose  $x \in E_n$  so that  $\pi(x) = \mathbf{y}$ , and let  $C_1, \alpha > 0$  be as in the hypothesis of the lemma. Then for all  $0 < \rho \leq 1$ ,  $0 < \beta \leq \rho^\gamma$ , and  $\mathcal{L} \in \mathcal{H}$ , if  $B = B(\mathbf{y}, \rho)$  then

$$\begin{aligned} \bar{\mu}(\mathcal{N}(\mathcal{L}, \beta\rho) \cap B \cap F_n) &\leq 2\nu_n(\mathcal{N}(\mathcal{L}, \beta\rho) \cap B) \\ &= 2\mu(\pi^{-1}(\mathcal{N}(\mathcal{L}, \beta\rho) \cap B) \cap E_n) \\ &\leq 2C_1 \beta^\alpha \mu(\pi^{-1}(B)) = 2C_1 \beta^\alpha \bar{\mu}(B). \end{aligned} \quad \square$$

**Notation 4.2.** In the sequel we will not distinguish between the sets  $S$  and  $\pi^{-1}(S)$ , so for example formula (4.1) would be written

$$\mu(\mathcal{N}(\mathcal{L}, \beta\rho) \cap B \cap E_n) \leq C_1 \beta^\alpha \mu(B).$$

We now state our general theorem about measures coming from “geometrically nice” coding maps.

**Theorem 4.3.** *Let  $A$  be a measurable space, let  $\pi : A^\mathbb{N} \rightarrow \mathbb{R}^d$  be a measurable map, and let  $\mu$  be a probability measure on  $A^\mathbb{N}$ . For each  $\omega \in A^*$ , we let  $\mu_\omega$  denote the conditional measure of  $\mu$  on the cylinder  $[\omega] = \{\tau \in A^\mathbb{N} : \tau \upharpoonright |\omega| = \omega\}$  (normalized to be a probability measure), and we fix a real number  $D_\omega \geq \text{Diam}(\text{Supp}(\mu_\omega))$ . Assume that  $D_\tau \leq D_\omega$  whenever  $\tau$  extends  $\omega$ . Fix  $\kappa > 0$ ,  $r \in \mathbb{N}$ , and a set*

$$(4.2) \quad G \subseteq \{\omega \in A^* : \mu_\omega(\{\tau \in A^\mathbb{N} : \text{Supp}(\mu_{\tau \upharpoonright |\omega|+r}) \cap \mathcal{N}(\mathcal{L}, \kappa D_\omega) = \emptyset\}) \geq \kappa \ \forall \mathcal{L} \in \mathcal{H}\}.$$

Assume that for  $\mu$ -a.e.  $\omega \in A^{\mathbb{N}}$ , the limits

$$(4.3) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \# \{i = 1, \dots, n : \omega \upharpoonright i \in G\}$$

$$(4.4) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log(1/D_{\omega \upharpoonright n})$$

exist and are positive. Then  $\pi_*[\mu]$  is quasi-decaying.

We will prove the theorem for the case  $r = 1$ ; the general case follows by replacing  $A$  by  $A^r$ .

The basic idea of the proof is as follows: Suppose that we are given a  $\mu$ -random word  $\tau \in A^{\mathbb{N}}$ , of which we know an initial segment  $\omega = \tau \upharpoonright n$ . We want to give an upper bound on the probability that  $\pi(\tau) \in \mathcal{N}(\mathcal{L}, \kappa\rho)$ , where  $\mathcal{L} \in \mathcal{H}$  and  $\rho > 0$ . Now reveal the letters of  $\tau$  in order. Each time a letter is revealed, there is a chance that the letter proves that  $\tau \notin \mathcal{N}(\mathcal{L}, \kappa\rho)$ . More precisely, if we are revealing the  $(i+1)$ st letter, and if  $\tau \upharpoonright i \in G$  and  $D_{\tau \upharpoonright i} \geq \rho$ , then the probability that  $\text{Supp}(\mu_{\tau \upharpoonright i+1})$  (which contains  $\tau$ ) is disjoint from  $\mathcal{N}(\mathcal{L}, \kappa\rho)$  is at least  $\kappa$ . So the probability that  $\tau \in \mathcal{N}(\mathcal{L}, \kappa\rho)$  is bounded above by an expression like  $(1 - \kappa)^k$ , where  $k$  is the number of  $i \geq n$  such that  $\tau \upharpoonright i \in G$  and  $D_{\tau \upharpoonright i} \geq \rho$ . The hypotheses (4.3)-(4.4) can be used to give a lower bound on  $k$ , which in turn gives an upper bound on probability which depends only on  $\omega$  and  $\rho$ .

We now proceed to make this idea rigorous:

*Proof of Theorem 4.3.* For each  $n \in \mathbb{N}$ ,  $\rho > 0$ , and  $k \in \mathbb{N}$  let

$$E(n, \rho, k) = \{\tau \in A^{\mathbb{N}} : \#\{i \geq n : D_{\tau \upharpoonright i} \geq \rho, \tau \upharpoonright i \in G\} \geq k\}.$$

**Claim 4.4.** For all  $\omega \in A^*$ ,  $\mathcal{L} \in \mathcal{H}$ ,  $\rho > 0$ , and  $k \in \mathbb{N}$ ,

$$(4.5) \quad \mu_{\omega}(\mathcal{N}(\mathcal{L}, \kappa\rho) \cap E(|\omega|, \rho, k)) \leq (1 - \kappa)^k.$$

*Proof.* For ease of exposition we assume that  $c = \inf_{\omega} \log(D_{\omega \upharpoonright |\omega|-1}/D_{\omega}) > 0$ , and we proceed by induction on  $\lfloor (1/c) \log(D_{\omega}/\rho) \rfloor$ . If  $\rho > D_{\omega}$ , then either  $E(|\omega|, \rho, k) = \emptyset$  or  $k = 0$ , and in either case (4.5) holds trivially. So assume that  $\rho \leq D_{\omega}$ . Write  $\ell = \lfloor \log_{\rho} D_{\omega} \rfloor$  (recall Convention 5). We have

$$\begin{aligned} & \mu_{\omega}(\mathcal{N}(\mathcal{L}, \kappa\rho) \cap E(|\omega|, \rho, k)) \\ &= \mu_{\omega}(\mathcal{N}(\mathcal{L}, \kappa\rho) \cap E(|\omega| + 1, \rho, k - \ell)) && (\text{since } \rho \leq D_{\omega}) \\ &= \int \mu_{\tau \upharpoonright |\omega|+1}(\mathcal{N}(\mathcal{L}, \kappa\rho) \cap E(|\omega| + 1, \rho, k - \ell)) \, d\mu_{\omega}(\tau) && (\text{conditional measures}) \\ &\leq \int (1 - \kappa)^{k-\ell} [\text{Supp}(\mu_{\tau \upharpoonright |\omega|+1}) \cap \mathcal{N}(\mathcal{L}, \kappa\rho) \neq \emptyset] \, d\mu_{\omega}(\tau) && (\text{induction hypothesis}) \\ &\leq (1 - \kappa)^{k-\ell} \mu_{\omega}(\{\tau \in A^{\mathbb{N}} : \text{Supp}(\mu_{\tau \upharpoonright |\omega|+1}) \cap \mathcal{N}(\mathcal{L}, \kappa D_{\omega}) \neq \emptyset\}) && (\text{since } \rho \leq D_{\omega}) \\ &\leq (1 - \kappa)^{k-\ell} (1 - \kappa)^{\ell} && ((4.2) \text{ and definition of } \ell) \\ &= (1 - \kappa)^k \end{aligned}$$

which completes the induction step.

The lemma can be proven without using the assumption  $\inf_{\omega} \log(D_{\omega \upharpoonright |\omega|-1}/D_{\omega}) > 0$  by using the martingale theorem instead of induction; the essential calculations are the same and we omit the details.  $\triangleleft$

Now let  $E \subseteq A^{\mathbb{N}}$  be a set on which the limits (4.3)-(4.4) converge uniformly. Fix  $\gamma > 0$ ,  $\mathbf{x} \in \mathbb{R}^d$ ,  $0 < \rho \leq 1$ ,  $0 < \beta \leq \rho^{\gamma}$ , and  $\mathcal{L} \in \mathcal{H}$ , and we will show that

$$(4.6) \quad \mu(\mathcal{N}(\mathcal{L}, \beta\rho) \cap B(\mathbf{x}, \rho) \cap E) \lesssim_{\times} \beta^{\alpha} \mu(B(\mathbf{x}, 2\rho))$$

for some  $\alpha > 0$  depending only on  $\gamma$ .

Consider the partition  $\mathcal{A}$  of  $A^{\mathbb{N}}$  consisting of all cylinders  $[\omega]$  ( $\omega \in A^*$ ) satisfying

$$(4.7) \quad D_{\omega} < \rho \leq D_{\omega \upharpoonright |\omega|-1}.$$

Fix such an  $\omega$ , let  $n = |\omega|$ , and let  $k = \lceil \log_{\lambda}(\beta) \rceil$ , where  $\lambda \in (0, 1)$  is small to be determined.

**Claim 4.5.** *If  $\beta$  is sufficiently small then*

$$[\omega] \cap E \subseteq E(|\omega|, \kappa^{-1}\beta\rho, k).$$

*Proof.* Fix  $\tau \in [\omega] \cap E$ . Let  $\delta = (1/2)\log_\rho(\beta) \geq \gamma/2$ , let  $\varepsilon = (\gamma/12) \wedge (1/5)$ , and let  $\ell = \lfloor \delta n \rfloor$ . By (4.7) and the definition of  $E$ , we have<sup>15</sup>

$$(4.8) \quad \#\{i = n+1, \dots, n+\ell : \tau \upharpoonright i \in G\} \gtrsim_+ g(\delta - \varepsilon)n$$

$$(4.9) \quad \rho < D_{\tau \upharpoonright n-1} \lesssim_\times \exp(-\chi(1 - \varepsilon)n),$$

$$(4.10) \quad D_{\tau \upharpoonright n+\ell} \gtrsim_\times \exp(-\chi(1 + \delta + \varepsilon)n)$$

$$(4.11) \quad \rho \geq D_{\tau \upharpoonright n} \gtrsim_\times \exp(-\chi(1 + \varepsilon)n),$$

where  $g, \chi > 0$  are the limits of (4.3) and (4.4), respectively. To show that  $\tau \in E(|\omega|, \kappa^{-1}\beta\rho, k)$ , it suffices to demonstrate the separate claims

$$(4.12) \quad D_{\tau \upharpoonright n+\ell} \geq \kappa^{-1}\beta\rho$$

$$(4.13) \quad \#\{i = n+1, \dots, n+\ell : \tau \upharpoonright i \in G\} \geq \lceil \log_\lambda(\beta) \rceil$$

under the assumption that  $\beta$  is sufficiently small. To prove (4.12), we observe that

$$\begin{aligned} \kappa\rho^{-1}D_{\tau \upharpoonright n+\ell} &\gtrsim_\times \exp(-\chi(\delta + 2\varepsilon)n) && \text{(by (4.9) and (4.10))} \\ &\gtrsim_\times \rho^{(\delta+2\varepsilon)/(1-\varepsilon)} && \text{(by (4.9))} \\ &\geq \rho^{(\delta+\gamma/6)/(1-1/5)} \geq \rho^{(4/3)\delta/(4/5)} = \beta^{5/6}. \end{aligned}$$

This implies that (4.12) holds for all sufficiently small  $\beta$ . To prove (4.13), we observe that

$$\begin{aligned} &\#\{i = n+1, \dots, n+\ell : \tau \upharpoonright i \in G\} \\ &\gtrsim_+ g(\delta - \varepsilon)n && \text{(by (4.8))} \\ &\gtrsim_+ \frac{g(\delta - \varepsilon)}{\chi(1 + \varepsilon)} \log(1/\rho) && \text{(by (4.11))} \\ &\geq \frac{g(\delta - \gamma/12)}{\chi(1 + 1/5)} \log(1/\rho) \geq \frac{(5/6)g}{(6/5)\chi} \delta \log(1/\rho) > \frac{g}{3\chi} \log(1/\beta). \end{aligned}$$

Choosing  $\lambda < \exp(-g/3\chi)$  gives (4.13) for all sufficiently small  $\beta$ .  $\triangleleft$

Combining Claims 4.4 and 4.5 yields (for  $\beta$  sufficiently small)

$$\mu_\omega(\mathcal{N}(\mathcal{L}, \beta\rho) \cap E) \leq (1 - \kappa)^k \leq \beta^\alpha,$$

where  $\alpha = \log_\lambda(1 - \kappa) > 0$ . On the other hand, by (4.7) we have  $\text{Diam}(\text{Supp}(\mu_\omega)) \leq D_\omega < \rho$ , so either  $\text{Supp}(\mu_\omega) \cap B(\mathbf{x}, \rho) = \emptyset$  or  $\text{Supp}(\mu_\omega) \subseteq B(\mathbf{x}, 2\rho)$ . Either way we have

$$\mu_\omega(\mathcal{N}(\mathcal{L}, \beta\rho) \cap B(\mathbf{x}, \rho) \cap E) \leq \beta^\alpha \mu_\omega(B(\mathbf{x}, 2\rho))$$

and integrating over all  $\omega$  yields (4.6). Since (4.6) is an asymptotic, it holds for all  $\beta$  rather than just for all sufficiently small  $\beta$ . Now [Part I, Lemma 3.2] shows that for  $\pi_*[\mu]$ -a.e.  $\mathbf{x} \in \mathbb{R}^d$ , we have

$$\mu(\mathcal{N}(\mathcal{L}, \beta\rho) \cap B(\mathbf{x}, \rho) \cap E) \leq \beta^{\alpha/2} \mu(B(\mathbf{x}, \rho)).$$

Since  $E$  was arbitrary subject to the condition that (4.3)-(4.4) converge uniformly, Egoroff's theorem and Lemma 4.1 show that  $\pi_*[\mu]$  is quasi-decaying.  $\square$

---

<sup>15</sup>In the following asymptotics, the implied constant may depend on  $\kappa$  and  $\gamma$  (and thus also on  $\varepsilon$ ), but not on other variables such as  $\delta$ .



**4.2. Definitions.** Before proving Theorem 1.14, we recall the definition of a conformal iterated function system (CIFS), its coding map, and the Gibbs measures of summable locally Hölder continuous functions.

**Definition 4.6** (Cf. [19, p.6-7]). Fix  $d \in \mathbb{N}$ . A collection of maps  $(u_a)_{a \in A}$  is called a *conformal iterated function system* on  $\mathbb{R}^d$  if:

1.  $A$  is a countable (finite or infinite) index set;
2.  $X \subseteq \mathbb{R}^d$  is a nonempty compact set which is equal to the closure of its interior;
3. For all  $a \in A$ ,  $u_a(X) \subseteq X$ ;
4. (Cone condition)

$$\inf_{\mathbf{x} \in X, r \in (0,1)} \frac{\lambda(X \cap B(\mathbf{x}, r))}{r^d} > 0,$$

where  $\lambda$  denotes Lebesgue measure on  $\mathbb{R}^d$ ;

5.  $V \subseteq \mathbb{R}^d$  is an open connected bounded set such that  $d(X, \mathbb{R}^d \setminus V) > 0$ ;
6. For each  $a \in A$ ,  $u_a$  is a conformal homeomorphism from  $V$  to an open subset of  $V$ ;
7. (Uniform contraction)  $\sup_{a \in A} \sup |u'_a| < 1$ , and if  $A$  is infinite,  $\lim_{a \in A} \sup |u'_a| = 0$ ;
8. (Bounded distortion property) For all  $n \in \mathbb{N}$ ,  $\omega \in A^n$ , and  $\mathbf{x}, \mathbf{y} \in V$ ,

$$(4.14) \quad |u'_\omega(\mathbf{x})| \asymp_\times |u'_\omega(\mathbf{y})|,$$

where

$$u_\omega = u_{\omega_1} \circ \cdots \circ u_{\omega_n}.$$

The CIFS is said to satisfy the *open set condition* if the collection  $(u_a(\text{Int}(X)))_{a \in A}$  is disjoint. It is said to satisfy the *strong separation condition* if the collection  $(u_a(X))_{a \in A}$  is disjoint.

Finally, the CIFS is said to be *irreducible* if there is no proper real-analytic submanifold  $M \subseteq X$  such that  $u_a(M) \subseteq M$  for all  $a \in A$ . (Equivalently, the limit set of the CIFS (defined below) is not contained in any proper real-analytic submanifold of  $\mathbb{R}^d$ .)

In the remainder of this section, we fix a CIFS  $(u_a)_{a \in A}$  and corresponding sets  $X, V \subseteq \mathbb{R}^d$ .

**Definition 4.7.** The *coding map* of the CIFS  $(u_a)_{a \in A}$  is the map  $\pi : A^\mathbb{N} \rightarrow \mathbb{R}^d$  defined by the formula

$$\pi(\omega) = \lim_{n \rightarrow \infty} u_{\omega|_n}(x_0),$$

where  $x_0 \in X$  is an arbitrary point. By the Uniform Contraction hypothesis,  $\pi(\omega)$  exists and is independent of the choice of  $x_0$ . The *limit set* of the CIFS is the image of  $A^\mathbb{N}$  under the coding map, i.e.  $K = \pi(A^\mathbb{N})$ .

Note that by the Uniform Contraction hypothesis, the coding map is always Hölder continuous, assuming that the metric on  $A^\mathbb{N}$  is given by the formula

$$d(\omega, \tau) = \lambda^{|\omega \wedge \tau|},$$

where  $\lambda \in (0, 1)$  and  $\omega \wedge \tau$  is the longest word which is an initial segment of both  $\omega$  and  $\tau$ .

**Definition 4.8** ([20, §2]). A function  $\phi : A^\mathbb{N} \rightarrow \mathbb{R}$  is called *locally Hölder continuous*<sup>16</sup> if there exist  $C, \alpha > 0$  such that for all  $\omega, \tau \in A^\mathbb{N}$  such that  $\omega_1 = \tau_1$ ,

$$|\phi(\omega) - \phi(\tau)| \leq C d^\alpha(\omega, \tau).$$

A locally Hölder continuous function  $\phi : A^\mathbb{N} \rightarrow \mathbb{R}$  is called *summable* if

$$\sum_{a \in A} \sup_{[a]} e^\phi < \infty.$$

A measure  $\mu$  on  $A^\mathbb{N}$  is said to be a *Gibbs state* for  $\phi$  if there exists  $P \in \mathbb{R}$  such that for all  $\omega \in A^*$  and  $\tau \in [\omega]$ ,

$$(4.15) \quad \mu([\omega]) \asymp_\times \exp \left( \sum_{j=0}^{|\omega|-1} \phi(\sigma^j \tau) - P|\omega| \right).$$

<sup>16</sup>In [20], this was just called “Hölder continuous”, but the terminology “locally Hölder continuous” is now standard, to distinguish it from the same definition with the requirement that  $\omega_1 = \tau_1$  removed.

The number  $P$  is called the *pressure* of the family  $\Phi$ .

**Theorem 4.9** (Special case of [20, Corollary 2.7.5(c)]). *If  $\phi : A^{\mathbb{N}} \rightarrow \mathbb{R}$  is a summable locally Hölder continuous function, then there exists a unique Gibbs measure  $\mu$  for  $\phi$  which is invariant and ergodic under the shift map. Any other Gibbs measure is coarsely asymptotic to  $\mu$ .*

**4.3. Proof of Theorem 1.14.** We may assume without loss of generality that  $\mu$  is the Gibbs measure which is invariant and ergodic under the shift map. For each  $\omega \in A^*$  let  $D_\omega = \text{Diam}(\pi([\omega]))$ . By the Bounded Distortion Property, for all  $\omega \in A^{\mathbb{N}}$  and  $n \in \mathbb{N}$  we have

$$D_{\omega|_n} \asymp_{\times} \prod_{j=1}^n |u'_{\omega_j}(\pi \circ \sigma^j(\omega))|$$

and so combining (1.9) with the ergodic theorem gives (4.4) for  $\mu$ -a.e.  $\omega \in A^{\mathbb{N}}$ . Moreover, if  $G = A^*$ , then clearly (4.3) holds. So to complete the proof, we need to show that there exist  $\kappa > 0$  and  $r \in \mathbb{N}$  such that (4.2) holds with  $G = A^*$ .

Suppose not; then for all  $r, n \in \mathbb{N}$  there exist  $\omega^{(r,n)} \in A^*$  and  $\mathcal{L}_{r,n} \in \mathcal{H}$  such that

$$(4.16) \quad \mu_{\omega^{(r,n)}}(\{\tau \in A^{\mathbb{N}} : \pi([\tau \upharpoonright |\omega^{(r,n)}| + r]) \cap \mathcal{N}(\mathcal{L}_{r,n}, (1/n)D_{\omega^{(r,n)}})\}) < 1/n.$$

Fix  $\mathcal{L}_0 \in \mathcal{H}$ , and for each  $r, n$ , let  $f_{r,n} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a similarity such that  $f_{r,n}(\mathcal{L}_{r,n}) = \mathcal{L}_0$ ,  $|f'_{r,n}| = 1/D_{\omega^{(r,n)}}$ , and  $f_{r,n} \circ u_{\omega^{(r,n)}}(x_0)$  is bounded for  $x_0 \in X$  fixed. Then by the Bounded Distortion Property,  $(f_{r,n} \circ u_{\omega^{(r,n)}})_{r,n}$  is a normal family, and so we can find convergent subsequences

$$f_{r,n} \circ u_{\omega^{(r,n)}} \xrightarrow[n]{r} v_r \xrightarrow[r]{r} v.$$

Let  $K$  be the limit set of  $(u_a)_{a \in E}$ . By hypothesis,  $K$  is not contained in any proper real-analytic submanifold of  $\mathbb{R}^d$ ; in particular,  $K$  is not contained in  $v^{-1}(\mathcal{L}_0)$ . So choose  $\tau \in A^{\mathbb{N}}$  such that  $v \circ \pi(\tau) \notin \mathcal{L}_0$ . There exists  $r_0$  such that  $d(v \circ \pi([\tau \upharpoonright r_0]), \mathcal{L}_0) > 0$ . Since

$$d(v_r \circ \pi([\tau \upharpoonright r]), \mathcal{L}_0) \underset{r \geq r_0}{\geq} d(v_r \circ \pi([\tau \upharpoonright r_0]), \mathcal{L}_0) \xrightarrow[r]{r} d(v \circ \pi([\tau \upharpoonright r_0]), \mathcal{L}_0) > 0,$$

there exists  $r \geq r_0$  such that  $d(v_r \circ \pi([\tau \upharpoonright r]), \mathcal{L}_0) > 0$ . Since

$$\begin{aligned} \frac{1}{D_{\omega^{(r,n)}}} d(\pi([\omega^{(r,n)} * \tau \upharpoonright n]), \mathcal{L}_{r,n}) &= d(f_{r,n} \circ u_{\omega^{(r,n)}} \circ \pi([\tau \upharpoonright n]), \mathcal{L}_0) \\ &\xrightarrow[n]{r} d(v_r \circ \pi([\tau \upharpoonright r]), \mathcal{L}_0) > 0, \end{aligned}$$

for all sufficiently large  $n$  we have

$$d(\pi([\omega^{(r,n)} * \tau \upharpoonright r]), \mathcal{L}_{r,n}) \geq (1/n)D_{\omega^{(r,n)}}.$$

Combining with (4.16) shows that

$$\mu_{\omega^{(r,n)}}([\omega^{(r,n)} * \tau \upharpoonright r]) < 1/n.$$

But by (4.15),

$$\mu_{\omega^{(r,n)}}([\omega^{(r,n)} * \tau \upharpoonright r]) \asymp_{\tau} \exp \left( \sum_{j=0}^{r-1} \phi \circ \sigma^j(\tau) - Pr \right) \asymp_{\times, \tau, r} 1.$$

Since  $n$  can depend on  $\tau$  and  $r$ , this is a contradiction.

## REFERENCES

1. B. Barany, *On the Hausdorff dimension of a family of self-similar sets with complicated overlaps*, Fund. Math. **206** (2009), 49–59.
2. L. M. Barreira, Y. B. Pesin, and J. Schmeling, *Dimension and product structure of hyperbolic measures*, Ann. of Math. (2) **149** (1999), no. 3, 755–783.
3. A. Barrio Blaya and V. Jiménez López, *On the relations between positive Lyapunov exponents, positive entropy, and sensitivity for interval maps*, Discrete Contin. Dyn. Syst. **32** (2012), no. 2, 433–466.
4. W. Bergweiler and A. E. Eremenko, *Meromorphic functions with linearly distributed values and Julia sets of rational functions*, Proc. Amer. Math. Soc. **137** (2009), no. 7, 2329–2333.
5. R. Bowen, *Periodic points and measures for Axiom A diffeomorphisms*, Trans. Amer. Math. Soc. **154** (1971), 377–397.
6. R. Broderick, L. Fishman, D. Y. Kleinbock, A. Reich, and B. Weiss, *The set of badly approximable vectors is strongly  $C^1$  incompressible*, Math. Proc. Cambridge Philos. Soc. **153** (2012), no. 02, 319–339.
7. R. Broderick, L. Fishman, and D. S. Simmons, *Badly approximable systems of affine forms and incompressibility on fractals*, J. Number Theory **133** (2013), no. 7, 2186–2205.
8. T. Das, D. S. Simmons, and M. Urbański, *Geometry and dynamics in Gromov hyperbolic metric spaces: with an emphasis on non-proper settings*, <http://arxiv.org/abs/1409.2155>, preprint 2014.
9. ———, *Dimension rigidity in conformal structures*, <http://arxiv.org/abs/1504.01774>, preprint 2015.
10. A. E. Eremenko and S. J. van Strien, *Rational maps with real multipliers*, Trans. Amer. Math. Soc. **363** (2011), no. 12, 6453–6463.
11. L. Fishman, D. S. Simmons, and M. Urbański, *Diophantine properties of measures invariant with respect to the Gauss map*, J. Anal. Math. **122** (2014), 289–315.
12. F. Hofbauer, *Generic properties of invariant measures for continuous piecewise monotonic transformations*, Monatsh. Math. **106** (1988), no. 4, 301–312.
13. ———, *Local dimension for piecewise monotonic maps on the interval*, Ergodic Theory Dynam. Systems **15** (1995), no. 6, 1119–1142.
14. I. Inoquio-Renteria and J. Rivera-Letelier, *A characterization of hyperbolic potentials of rational maps*, Bull. Braz. Math. Soc. (N.S.) **43** (2012), no. 1, 99–127.
15. D. Y. Kleinbock, E. Lindenstrauss, and B. Weiss, *On fractal measures and Diophantine approximation*, Selecta Math. **10** (2004), 479–523.
16. D. Y. Kleinbock and G. A. Margulis, *Logarithm laws for flows on homogeneous spaces*, Invent. Math. **138** (1999), no. 3, 451–494.
17. D. Y. Kleinbock and B. Weiss, *Badly approximable vectors on fractals*, Israel J. Math. **149** (2005), 137–170.
18. R. D. Mauldin, T. Szarek, and M. Urbański, *Graph directed Markov systems on Hilbert spaces*, Math. Proc. Cambridge Philos. Soc. **147** (2009), 455–488.
19. R. D. Mauldin and M. Urbański, *Dimensions and measures in infinite iterated function systems*, Proc. London Math. Soc. (3) **73** (1996), no. 1, 105–154.
20. ———, *Graph directed Markov systems: Geometry and dynamics of limit sets*, Cambridge Tracts in Mathematics, vol. 148, Cambridge University Press, Cambridge, 2003.
21. S. J. Patterson, *The limit set of a Fuchsian group*, Acta Math. **136** (1976), no. 3–4, 241–273.
22. F. Paulin, M. Pollicott, and B. Schapira, *Equilibrium states in negative curvature*, <http://arxiv.org/abs/1211.6242>, preprint 2013.
23. M. Pollicott and K. Simon, *The Hausdorff dimension of  $\lambda$ -expansions with deleted digits*, Trans. Amer. Math. Soc. **347** (1995), no. 3, 967–983.
24. A. D. Pollington and S. L. Velani, *Metric Diophantine approximation and “absolutely friendly” measures*, Selecta Math. **11** (2005), 297–307.
25. F. Przytycki and M. Urbański, *Conformal fractals: ergodic theory methods*, London Mathematical Society Lecture Note Series, 371, Cambridge University Press, Cambridge, 2010.
26. J. Rivera-Letelier, *The maximal entropy measure detects non-uniform hyperbolicity*, Math. Res. Lett. **17** (2010), no. 5, 851–866.
27. K. Sigmund, *Generic properties of invariant measures for Axiom A diffeomorphisms*, Invent. Math. **11** (1970), 99–109.
28. D. S. Simmons, *Conditional measures and conditional expectation; Rohlin’s disintegration theorem*, Discrete Contin. Dyn. Syst. **32** (2012), no. 7, 2565–2582.
29. K. Simon and B. Solomyak, *On the dimension of self-similar sets*, Fractals **10** (2002), no. 1, 59–65.
30. B. O. Stratmann and M. Urbański, *Diophantine extremality of the Patterson measure*, Math. Proc. Cambridge Philos. Soc. **140** (2006), 297–304.
31. B. O. Stratmann and S. L. Velani, *The Patterson measure for geometrically finite groups with parabolic elements, new and old*, Proc. London Math. Soc. (3) **71** (1995), no. 1, 197–220.
32. D. P. Sullivan, *The density at infinity of a discrete group of hyperbolic motions*, Inst. Hautes Études Sci. Publ. Math. **50** (1979), 171–202.
33. ———, *Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups*, Acta Math. **153** (1984), no. 3–4, 259–277.

34. M. Szostakiewicz, M. Urbański, and A. Zdunik, *Fine inducing and equilibrium measures for rational functions of the Riemann sphere*, preprint 2011.
35. M. Urbański, *Diophantine approximation and self-conformal measures*, J. Number Theory **110** (2005), 219–235.
36. ———, *Diophantine approximation for conformal measures of one-dimensional iterated function systems*, Compos. Math. **141** (2005), no. 4, 869–886.

UNIVERSITY OF WISCONSIN – LA CROSSE, DEPARTMENT OF MATHEMATICS, 1725 STATE STREET, LA CROSSE, WI 54601, USA

*E-mail address:* `tdas@uwlax.edu`

*URL:* <https://sites.google.com/a/uwlax.edu/tdas/>

UNIVERSITY OF NORTH TEXAS, DEPARTMENT OF MATHEMATICS, 1155 UNION CIRCLE #311430, DENTON, TX 76203-5017, USA

*E-mail address:* `lior.fishman@unt.edu`

UNIVERSITY OF YORK, DEPARTMENT OF MATHEMATICS, HESLINGTON, YORK YO10 5DD, UK

*E-mail address:* `simmons.465@osu.edu`

*URL:* <https://sites.google.com/site/davidsimmonsmath/>

UNIVERSITY OF NORTH TEXAS, DEPARTMENT OF MATHEMATICS, 1155 UNION CIRCLE #311430, DENTON, TX 76203-5017, USA

*E-mail address:* `urbanski@unt.edu`

*URL:* <http://www.urbanskimath.com/>